GRADUATE SCHOOL
U. S. DEPARTMENT OF AGRICULTURE
SEMINAR ON AUTOMATIC DATA PROCESSING
FOR FEDERAL EXECUTIVES

Presentation by Mr. V. J. Fogarty
U. S. Dept. of Agriculture Graduate School
Washington, D. C.

The opportunity to present this session of the graduate school's seminar on automatic data processing is a pleasant assignment and something I've been looking forward to.

I like to talk about EDPM because the possibilities it offers are so tremendous. This is really a big subject and it reminds me of the familiar fable about the four blind men who came upon an elephant and, after investigating the strange creature, tried to describe it. You'll recall that the first man seized on one of the beast's legs and reported that the elephant resembled a tree trunk.

Another of them found himself up against the creature's side and described an elephant as much like a rough and weatherbeaten wall. A third chanced to grasp the animal's trunk. . .to him, the creature resembled a heavy hawser, or a great snake or serpent of some kind.

The fourth man grasped, I believe, the elephant's ear and discovered that the animal was very like a bird with heavy, flexible wings. Or it may be. . .I'm not too sure of my recollection here. . .that the fourth took hold of another part of the creature and received a distinctly different impression. It doesn't really matter. The point is clear enough. . .that our reactions to a thing seem to depend on what part of its anatomy we grasp.

Now I'm not going to try to describe in detail, the whole elephant, but rather just that part of our EDP program concerned with pre-installation planning after it was decided that we would install a machine. Then we can discuss very briefly what has been going on since our first machine was installed in 1956.

Before we get into a detailed discussion of our EDP program I believe we should spend a few minutes discussing the Ships Parts Control Center and what its mission is.

The Ships Parts Control Center--SPCC--is the Navy's Supply-Demand Control Point, or inventory manager, for ships repair parts. Its mission is to insure a proper balance between the supply of and demand for these parts. The SPCC is assigned the inventory control responsibility for such items as: (1) equipments and repair parts for Internal Combustion Engines; (2) Gyrocompass, Main Propulsion and Navigation equipment; and (3) many other electrical, hull and machinery equipments or components.

SPCC is the catch-all Supply-Demand Control Point. If the components or equipments installed aboard ship are not pure electronics, pure ordnance, pure aviation, peculiar to submarines, or commonly used everyday general stores items, the repair parts support responsibility is usually assigned to the SPCC.

Technically, SPCC comes under the control of the Bureau of Ships and under the management control of the Bureau of Supplies and Accounts. It maintains a centralized inventory control on 135,000 items. At one time the SPCC maintained

control of more than 300,000 items in store, mainly due to the various types of equipment it is required to support.

But many of these items were added to SPCC's inventory list during wartime and some were based on the recommendations of manufacturers that certain repair parts should be carried in the supply system. A lot of this material did not "move", and has been disposed of. At the present time, the SPCC controls an inventory valued at approximately $450 million.

It has been estimated that the SPCC could be called upon to supply, either from stock or by purchase, a range of from one to three million items of repair parts. Because it is impractical to stock all of these items in the supply system, centralized inventory control is maintained over only 135,000 items.

Our investigation for the possible installation of an EDPM started in 1952, at least in the dreaming stage, and installation of card-to-tape equipment was made in May 1956, followed by the central processing unit in August 1956.

I have prepared a pre-installation planning chart which I will use as a guide for our discussion this morning. This chart covers a period of 15 months prior to installation of the machine. It is considered that about fifteen months is required in order to get ready for a machine. There's a lot to be done during this 15 months and it will be helpful for you to have an outline to follow.

ORGANIZATION - 15 MONTHS

The first action required is to insure satisfactory installation of the data processing equipment.

EXECUTIVE COMMITTEE

The extensive changes in organizational responsibility, personnel, systems, and methods, and the large potential for dollar and personnel savings that result from EDPM use requires that top management direct, coordinate, and approve the EDPM program. To insure this direction, it is recommended that an EDP Executive Committee be established and charged with the full responsibility of progress in pre-installation planning and will make necessary management decisions affecting policy, procedures, personnel, etc. This committee will advise and evaluate all phases of the operations for top management.

PROJECT DIRECTOR

Reports to the Executive Committee and is responsible for supervision of all details of the installation program including coordination, assignment, scheduling, and follow-up.

ANALYSIS AND PROGRAMMING STAFF

This staff will be assigned to specific duties by the Project Director: (1) it will make studies in areas selected; (2) design procedures for data handling; (3) write programs; (4) prepare test data for the various machine runs; (5) advise the Project Director of its progress; and (6) report on all phases requiring management decision.

control of more than 300,000 items in stock, mainly due to the variety of types of equipment it is required to support.

But many of these items were added to SPCC's inventory list during wartime and some were based on the recommendations of manufacturers that certain repair parts should be carried in the supply system. A lot of this material did not move, and has been disposed of. At the present time, the SPCC controls an inventory valued at approximately $420 million.

It has been estimated that the SPCC could be called upon to supply, either from stock or by purchase, a range of from one to three million items of repair parts. Because it is impractical to stock all of these items in the supply system, centralized inventory control is maintained over only 313,000 items.

Our investigation for the possible installation of an EDPM started in 1956, at least in the planning stage, but installation of card-to-tape equipment was made in May 1957, followed by the computer processing unit in August 1958.

I have prepared a pre-installation planning chart which I will use as a guide for our discussion this morning. This chart covers a period of 15 months prior to installation of the machine. It is considered that about fifteen months is required in order to get ready for a machine. There's a lot to be done during this 15 months and it will be helpful for you to have something to follow.

ORGANIZATION - 15 MONTHS

The first action required is to insure satisfactory installation of the data processing equipment.

EXECUTIVE COMMITTEE

The extensive changes in organizational responsibility, personnel, systems and methods, and the large potential for major improvement savings that result from EDPM use requires that top management direct, coordinate, and support the EDPM program. To insure this situation, it is recommended that an EDP Executive Committee be established and charged with the full responsibility of program of preinstallation planning and will make decisions and management decisions affecting policy, procedures, personnel, etc. This committee will review and evaluate all phases of the operation for top management.

PROJECT DIRECTOR

Reports to the Executive Committee and is responsible for supervision of all details of the installation program including coordinating, management, scheduling, and follow-up.

ANALYSIS AND PROGRAMMING STAFF

This staff will be assigned to specific duties by the Project Director. (1) It will make studies in areas selected, (2) design procedures for data handling, (3) write programs, (4) prepare test data for the various machine runs, (5) advise the Project Director of its progress, and (6) report on all phases regarding management decision.

At Mechanicsburg we had no formal directive for initiating the original feasibility study. The EDPM study was initiated in 1952 by key personnel of the Machine Records Division. Our initial report was approved by the Commanding Officer and forwarded to the Bureau of Supplies and Accounts by letter dated 6 April 1954. The Bureau approved the SPCC recommendations for installation of a machine and placed a Letter of Intent with IBM for delivery of a 702. We never did install the 702; however, because IBM only made a small number of these machines. If my memory serves me correctly they only built 18 of the 702 machines. We installed the first 705 in the Navy.

Our Executive Committee was established in June 1954. Members were selected based upon their intimate knowledge of specific areas for application of EDPM techniques. The committee was comprised of subject matter specialists and top level supervisory personnel who would ultimately be responsible for implementing committee recommendations. Commander H. P. Mills, Machine Records Division Director, was selected as Chairman by the Commanding Officer, in view of his previous data processing experience and the part he played in developing the initial feasibility study.

ADMINISTRATION - EDUCATION

Your Project Director must establish administrative plans and objectives for the groups involved. He must set up lines of communications and reporting procedures and provide for follow-up on problems and progress. In order to assist each of the administrative groups in the proper evaluation and execution of its respective functions, arrangements must be made to provide sufficient data processing machine training. This education should follow the general pattern indicated.

1. EXECUTIVE COMMITTEE

Each member of this committee should attend a Data Processing Executive Class. This training should also be supplemented by orientation seminars and demonstration visits where possible.

2. PROJECT DIRECTOR

Should attend Data Processing Programming Class. He should also be included in any of the seminars or demonstrations conducted for the Executive Committee.

3. ANALYSIS AND PROGRAMMING STAFF

All members of this staff should have a thorough knowledge of programming. Each member of the group should complete a programming class.

SCHEDULING

The Project Director, at this time, should -

Discuss with the Executive Committee the projected delivery schedule. He should emphasize the time required and the work to be accomplished during the pre-installation period and insure that both the customer and the manufacturer will be ready for machine installation on the date of scheduled delivery.

I can assure that if you are able to accomplish the objectives set forth, so far, 15 months prior to installation of your machine--you are in excellent shape---

KEEP UP THE GOOD WORK!

Throughout the course of this fifteen months period your Project Director is going to be "involved" in many, many things and you will need someone who won't lose sight of the "big picture". Speaking of "getting involved" in things reminds me of the announcement of a professor's new automatic computer and his wife's new baby which appeared almost simultaneously in the local newspaper. Upon being congratulated on "This proud event in the family", the professor naturally thought first of the achievement that had cost him the greater effort. "Thank you" he replied modestly, "but I couldn't have done it without the help of two graduate students".

You may not agree with me when I say that the professor lost sight of the big picture. So, I say, be very discriminating and selective in deciding who your Project Director shall be.

OBJECTIVES - 14 MONTHS
GENERAL FLOW CHART
POLICY CHANGES
PERSONNEL

OBJECTIVES

At this time the Executive Committee should be preparing a broad definition of the objectives of the program, the selection of applications and their priority, and also the target dates for completion of various levels of activity. You should consider the following points in application selection:

1. Desire for rapid cut-over operations.
2. Desire for early savings.
3. Possibility of centralization.

At Mechanicsburg, inventory control (redistribution, reallocation and procurement) and the maintenance and dissemination of technical engineering records were the starting applications.

We could foresee the day when many significant improvements could be achieved if high speed data processing at SPCC would become a reality. For example, at the time we submitted our request to the Bureau for installation of an EDPM, the SPCC was receiving quarterly stock status reports on each item of ships parts carried at our reporting activities. These items were reported on whether any transactions had occurred or not. The information was in summary form and the volume of cards was so great that more than a month went by before our EAM machines produced the consolidated report for review by our Stock Control Analysts.

We knew that a great many things were happening to our material, but we did not have the means to always take definite, and timely action. That is one of the primary reasons why we wanted to obtain an Electronic Data Processing Machine.

We wanted to change our reporting system from the quarterly, summary type of field reporting to a transaction reporting system. By the use of high speed magnetic tape machines, we would be able to process our reports rapidly and take timely action. Our plan for changing the reporting system was approved and implemented in October 1956.

With transaction reporting and the use of management by exception techniques, all items are reviewed periodically, but only those items requiring action are

printed out for the Stock Control Division. In doing this, the machine makes many decisions. This permits Stock Control personnel to devote more time to those items requiring attention. Also, all the known facts are associated with the item. This was not possible with the old punch card system. We have combined into our Master Tape records all of the information previously contained in ten different punch card files. All of this information is now consolidated on one piece of paper - the CSSR.

The items, which need attention, are processed by the machine with resulting automatic actions proposed to the Stock Control Analyst.

Another important factor is that the machine is always consistent but many times people working the very same instructions come up with conflicting answers. The machine processes an item the same way all the time because we have instructed it to do so.

MAINTENANCE OF CATALOG AND TECHNICAL RECORD FILES

Another really big application for the EDPM was automatic record maintenance.

Our Technical Records files (about 10 million cards) were converted to tape in September 1956 and have been maintained on a monthly basis since that time. Maintenance every month was not feasible economically with conventional EAM equipment and, under the old system, approximately six months were required to complete a full processing cycle. The 705 has made it possible to maintain the file on a current basis.

I might say that we were pretty successful in installing our machine--but I really think it was the result of establishing our objectives and planning to meet these objectives. We stressed the importance of the organization behind the machine. The ability of those in charge to keep their feet on the ground and not float away on cloud 9 was also a valuable attribute.

GENERAL FLOW CHART

The Project Director and the Analysis and Programming Staff should prepare a general flow chart of the objectives prescribed by the Executive Committee. Emphasis should be placed on responsibility of the Analysis and Programming Staff and the importance of a schedule for project completion.

The general flow charts prepared at the time of the feasibility study could be updated for this purpose and once completed, should be reviewed by the Project Director and the Executive Committee. Approval should be obtained from all responsible executives. Flow charts show major machine operations and do not intend to display the machine operation in detail.

A narrative explanation should accompany the flow charts to explain each major operation in detail. The narrative must be explicit and correct as machine operation will be tailored from it. Consolidation of present files should be noted at this point. The programmer may elect to make a further consolidation or breakout of files to facilitate machine operation, but in general the decisions made now will hold true.

POLICY CHANGES

Policy changes resulting from the general flow chart; or changes that affect personnel, procedures and public relations should all be discussed with appropriate personnel.

See to it that necessary steps are taken to initiate these changes as they occur.

PERSONNEL

The Project Director should work together with the Personnel Department on manpower requirements. He should give consideration to the following points:

1. Type of personnel, whether they be:

a. Supervisory level - could come from the Analysis and Programming Staff. If they don't, they should attend programming school.

b. Tape Librarian - could come from Analysis and Programming Staff. If they don't, they should attend programming school.

c. Console Operators - should have console and machine operation training. Should be able to read and understand a program.

d. Machine Operators (punched card and EDP) - should be given normal training for EAM equipment but Computer Operators should be given training in EDP operation and card and tape handling. We have even gone so far as to have our GS-5 Computer Operators go to programming school after they have been on the job about six months. They like to have the opportunity and we feel it pays big dividends.

PERSONNEL SELECTION

We had a real problem in the personnel area in staffing the positions of responsibility. We were very careful in selecting our EDP Staff. Our selection panels spent a great deal of time and effort reviewing past experience and the work records of the candidates for our positions. For the persons to be interviewed we prepared our questions in advance so as to be certain to cover everything the panel wanted to obtain additional information on.

It was hard for our employees to meet the minimum Civil Service requirements for the Computer Programmer positions. In order to partially overcome this, we entered into a promotion and training agreement with the third regional office of the Civil Service Commission so that our employees might qualify for some of these jobs. For example, if an employee successfully completed the training program, he would qualify as a GS-9 Computer Programmer, obtaining the required specialized experience necessary in only nine months of training. In order to help make sure that the employees given the training would successfully complete the course, eighty-nine applicants were tested for the first eight programmer positions. We interviewed the top twenty-two applicants, and selected those who were considered as having the best overall knowledge of the Ships Parts Supply System, or at least had a thorough knowledge or familiarity with one phase or area of the system--such as Stock Control, Technical Records, Catalog, etc. By carefully selecting our candidates, we were confident that they would successfully complete the training program, and would, after they had finished the training program, be good employees for the new organization. We hoped we would not be just a training ground for industry. By being up in the sticks at Mechanicsburg, and by selecting candidates having some ground roots to the area, we hoped that we would not be faced with a big turnover problem.

It's been our practice in our interviews to select applicants who have established permanent residence in the area; people with a defined motive for applying for data processing work and who display the proper attitude towards a career in Federal Service.

RESTUDY SELECTED APPLICATIONS - 13 MONTHS
ORIENTATION

RESTUDY SELECTED APPLICATIONS

The Project Director and the Analysis and Programming Staff should make a detailed survey of the selected applications. This survey should emphasize the following points:

1. Observe objectives and results rather than the techniques used.

2. Take note of "cause and effect" areas relating to results in the base applications under study.

3. Pay special attention to exceptions. Can be very troublesome, must get decision rapidly.

4. Note existing due-in and due-out schedules. What is effect of EDP?

5. Analyze present manpower requirements. Increases as well as decreases.

6. Consider all costs of present operations

7. Rely on the normal survey techniques in the examination of:
   a. Source Documents.
   b. Files.
   c. Reports.
   d. Procedures.

ORIENTATION

The Project Director should coordinate with operating officials in the preparation and conduct of orientation programs for the personnel in affected areas.

The personnel situation that exists when an EDPM is to be installed can be a very touchy one and I'd like to tell you how we coped with it, but I believe it appropriate to discuss the entire personnel situation rather than just that portion concerning the EDPM operation. The Command at SPCC displayed a real concern for employee morale. The Command placed sincere emphasis on the need to keep employees informed of the progress of EDP planning. Naturally, all employees were wondering how the machine would effect their jobs and I believe a great deal of our success can be attributed to the fact that the employees were kept informed. After the machine was installed all employees of the Ships Parts Control Center were extended an invitation to visit the Data Processing Center and see the 705 machine in operation. The personnel were scheduled in groups of thirty-five at about twenty minute intervals. The visit began in the EDP "viewing room" where each piece of equipment was discussed. Interesting statistics and the speed of operations were highlighted and this was followed by a tour through the EDPM room itself.

Throughout the twenty minute visit, the underlying theme was to convey to these people that they would, either directly or indirectly, have some part in preparing information that would become an input to the machine or have some part in working with output information that had been processed by the machine. It was made clear that the machine could quickly perform their routine clerical operations permitting them more time for the important aspects requiring evaluation or decisions.

The effort we put forth in keeping the employees apprised of the EDPM program paid big dividends and when it came time to reduce personnel, it was done without anyone getting hurt. The EAM tabulating equipment operators positions were the ones most effected and most of these people found jobs as clerks or worked into the EDPM program. A number of employees transferred to another Naval installation at Mechanicsburg and still others obtained jobs in industry. The significant thing about all this is that an informed group of people cooperated fully in this endeavor. Everyone seemed to consider the EDPM operation as another forward step of progress.

We had other training to do also. We had to indoctrinate key management personnel. To acquaint top level supervisors with the operations of high speed data processing equipment military and key civilian personnel were progressively scheduled to attend a one week Executive Course conducted by IBM. We went "all out" on this type of training and even had several classes conducted right at Mechanicsburg. By the time the 705 was installed over 150 people had been afforded this type of training. You might even say we "brain washed" top supervisory personnel because we wanted their support.

SITE PLANNING - 12 MONTHS
SPACE REQUIREMENTS
POWER REQUIREMENTS
AIR CONDITIONING
CONSTRUCTION PLANS
CONTINUE STUDY

SPACE REQUIREMENTS

The Project Director and Executive Committee must make the initial determination as to the best location for the machine. Usually there are only one or two logical places to locate the machine. Many of the EDP machines require a controlled climate in which to operate, and thus a building which is already air conditioned will result in cheaper site costs.

The Project Director should request early assistance from Public Works Department personnel. In relation to size, consider the components in order, the tape file storage requirements, and the needs of the customer engineering group.

Prepare a preliminary layout. Consider again the points mentioned above as well as the requirements in the physical planning pre-installation manual provided by the EDP manufacturer. Take definite steps toward selection of a contractor.

POWER REQUIREMENTS

The Project Director, Physical Planning Engineer, and the Building Engineers should review the power requirements as set forth in the manufacturer's physical planning pre-installation manual. Consider all phases of power requirements including the possibility of both data-processing machines and punched-card equipment.

AIR CONDITIONING

It may be wise to use an outside consultant for A/C advice. Make sure that all responsible persons are in complete agreement as to your requirements. Stress need for control over conditions required by the physical planning pre-installation manual.

CONSTRUCTION PLANS

The Project Director and Physical Planning Engineer should insure that definite steps will be taken towards completion of the necessary construction work, selection of a contractor, and acquisition of necessary equipment. Make arrangements for a meeting to review progress of construction planning.

This Project Director is a busy boy, isn't he: Believe it, or not, the site construction problem proved to be very difficult and time consuming for us. IBM suggested that we start real early on site construction plans and we are most thankful that we heeded their advice. Our records show that as early as December 1954, some 19 months before our machine was to arrive we reported to the Bureau of Supplies and Accounts:

"Tentative plans for the physical installation of the electronic data processing machine have been developed. The Public Works Officer of the Naval Supply Depot, Mechanicsburg, is submitting a letter of request to the Bureau for approval of an architectural and engineering contract prior to receipt of funds from the Shore Station Development Project for the installation. It is believed that such a contract would be warranted because of the technicality of the air-conditioning and humidity controls necessary for this type of equipment."

BLOCK DIAGRAMMING - 11 MONTHS

COMPLETION OF RESTUDY

The Project Director and the Analysis and Planning Staff should complete the restudy of application areas during this month. Since the block diagramming is to be started in this period, manpower of the Analysis and Programming Staff must be diverted to this work.

BLOCK DIAGRAMMING

Block Diagrams are detailed machine operation charts. These charts are developed by the programmer and require a very detailed knowledge of the machine. The actual coding of the instructions will be done directly from these charts.

The Analysis and Programming Staff should begin preparing a block diagram for each data processing machine run. Since this is a graphic representation of the logical development of the problem, consideration should be given to the following points:

1. Control of data before the machine pass.
2. Type of input.
3. Programmed controls.
4. File maintenance
5. Program checking and error-correction routines.
6. Type of output.
7. Control of data after data processing machine pass.

BLOCK DIAGRAM REVIEW

Errors in the problem definition or logical solution can be costly and time consuming to correct. It is important that a review of the block diagrams are conducted as they are completed.

The chart I have prepared suggests that block diagramming be completed over a period of six months, but it may actually take a little longer because coding

and block diagramming are somewhat closely related. Estimated 75% of job is flow charting and block diagramming; 25% coding.

Our historical records tell us that we were doing our block diagramming about 15 months prior to installation of the machine for our stock control application. The following is a quotation from a report we submitted to our management Bureau in May 1955.

"During the months of March and April, preparation of detailed block diagrams for the stock control problem was undertaken. The stock control application has been divided into three segments as follows:

1. Daily processing for updating the Perpetual Inventory Records.
2. Preparation of shipment orders, redistributions, reallocations, procurements, etc.
3. Preparation of inventory analysis, budget statistics, and processing permanent changes to the perpetual records caused by stock record change notice actions.

It is considered that progress in this area has been satisfactory. The Stock Control Division has provided detailed criteria for use by the machine in making procurement recommendations, redistributions, etc., for the fast moving items. Detailed criteria for items other than the fast movers are presently being developed by the Stock Control Division. It is expected that these guidelines will be completely finalized within the next three to four weeks. The actual writing of program instructions for preparation of inventory analysis reports is already in process."

PROGRAMMING - 10 MONTHS
(CONTINUE BLOCK DIAGRAMMING AND REVIEW)
DISCUSS SHIFT USAGE

CONTINUE BLOCK DIAGRAM

The Project Director should insure that block diagramming continues on schedule.

CONTINUE BLOCK DIAGRAM REVIEW

Check on progress of the review of block diagramming. See that the review is keeping pace with the block diagram preparation.

PROGRAMMING

The Project Director should assist the Analysis and Programming Staff in writing programs from the block diagrams. Consider the following points in the organization of this work:

1. Program teams. Number of programmers to a team, area to be handled by a team, number of program steps within the area, and scheduled completion dates. Teams work well especially when they write their first program. Can help each other.

2. Utility routines. What routines are required? Sorting routines, print memory, End-of-File, etc. What is available from manufacturer of equipment. Many programs are available from IBM.

3. Check Point and Restart procedures.

SHIFT USAGE

A meeting with the customer engineering manager should be set up to review plans for customer engineering coverage. Establish the number of shifts required and the components to be used on each shift. Review the extra shift policy. Review schedules for diagnostic routines and bias checks.

PROGRAM TESTING - 9 MONTHS
(CONTINUE BLOCK DIAGRAMMING)
SPECIAL REQUIREMENTS

BLOCK DIAGRAMMING

The Project Director should ascertain that all block diagramming operations continue during this period and that they will be completed about five months before installation of the machine.

REVIEW OF BLOCK DIAGRAMMING

During the month the review of the block diagrams should continue until completed. Follow closely to see if additional manpower will be required to complete this work on schedule.

CONTINUE CODING

The Project Director should follow up on coding progress. As block diagramming is completed more effort should be shifted to the programming phases.

PROGRAM TESTING - Anxious to Test a Program

Prior to the arrival of the machine, testing at one of the test centers should take place. This is necessary in order to have the programs checked before arrival of the machine.

An early test session is recommended to familiarize the programmers with the machine. It is also a boost to morale as many of the programmers have never seen a machine. The first test session also serves to tell whether the programmers are on the right track in their machine logic. The first test session should be held as soon as there is sufficient programming completed to make the time worthwhile.

Test schedules for the rest of the test time available should be worked out well in advance. The last test session should be completed about one month before machine installation.

Only completed programs should be taken to the test sessions. These programs should have been well desk checked and any revisions should be made before the test session. There is not time to make any major changes in a program while at the test center.

The Project Director should consider special requirements for the test sessions such as security clearance and visits by executives during testing operations. Programs that exceed the capacity of the testing machines.

Again quoting from our historical records I can tell you that -

"On 7 and 8 November 1955, the Ships Parts Control Center completed the first check out period on the IBM Type 702 Electronic Data Processing machine at

the IBM Data Processing Center, New York City. The program to build perpetual inventory records for the stock control application was thoroughly checked. This problem required eleven hundred forty (1140) program steps and was processed without error the very first time SPCC representatives were given machine time. This command is proud of the personnel who contributed to this remarkable feat and is pleased to be able to submit such information to the Bureau of Supplies and Accounts. Approximately seven thousand (7,000) program steps have been written for the inventory control application. It has been estimated that this entire area of work will require from 10,000 to 12,000 steps. It is planned to test this program at the Data Processing Center in New York in March.

We later reported that during the months of April and May 1956, visits were made to the IBM Data Processing Center, New York City, for program testing purposes. Twelve (12) machine runs involving 18,445 program instructions were successfully tested."

SPECIAL REQUIREMENTS

The Project Director should check with the Analysis and Programming Staff on the details of any special requirements and/or devices. Watch for the necessity of delivery schedule tie-in for punched-card equipment. I believe I mentioned that we received our card to tape equipment in May 1956. We got this equipment about four months prior to the central processing unit so that we could get our cards converted to magnetic tape well in advance of installation of the main frame.

It became apparent to us in November 1955 we had a need for special requirements. The following is a quotation from a report to our management Bureau:

"The recent announcement by the IBM Corporation of the availability, after October 1956, of twenty thousand additional positions of magnetic core storage on the 705 machine was received with extreme interest. In the inventory control area the use of a central processing unit with 40,000 positions of magnetic core storage would eliminate the necessity of two machine runs each week and four machine runs each quarter. This is the equivalent of seven hours machine operating time each week and would provide additional machine time for the allowance list program, as well as for the conversion of additional programs to EDPM processing in the future. Because of the long lead time involved, it is recommended that reference (b) be amended to include 20,000 additional positions of magnetic core storage for installation as soon as possible. The increased rental amounts to $2500.00 per month. While the SPCC is not certain at this time, it is nevertheless believed that the additional storage will be sufficient to obviate the need for the magnetic drum. However, it is felt that the best interests of the Navy can be served by having the magnetic drum also remain on order until the requirements of the Simplified Allowance List Program are fully defined. In this regard, the Ships Parts Control Center recently completed a test on the preparation of thirty of the new Allowance Parts Lists for the Bureau of Ships to further develop this program."

TIME AND COST ANALYSIS - 8 MONTHS
CONVERSION REQUIREMENTS
(CONTINUE CODING)

TIME AND COST ANALYSIS

The Project Director and the Analysis and Programming Staff should establish time and cost factors to insure that the equipment will do the job and will

achieve the results you desire. These factors are the major elements in the justification of the system. Consideration should be given to the components on order; both data processing machines and punched-card equipment, as well as to the cost figures discussed during the survey. It's most important to keep these data available for later management review people.

I'm sure you all know where the management review people come from. They are people who would have been poets, historians or biographers, etc., if they could: they have tried their talents at one or the other, and have failed; therefore, they become reviewers. Or as Disraili has said; "It's much easier to be critical than to be correct."

Seriously, you should review the cost and time figures with the Executive Committee. Obtain approval of all factors. Insure that action is taken when required; e.g. additional components to be ordered or released.

CONVERSION REQUIREMENTS

The Project Director should establish schedules for conversion operations. The following points should be considered to see that steps are taken to meet these requirements:

1. Procedures and controls
2. Component requirements
3. Tape usage
4. Customer Engineering coverage
5. Power, space and air conditioning
6. Delivery schedules

THERE ARE TWO MAIN METHODS IN USE FOR CONVERSION:

1. "One Time" conversion will gather all the data needed to start the operation at once. It will all be processed through the conversion programs at the same time. Once that has been completed, regular operations will start. This has the advantage of being very straightforward and eliminates the operation of two jobs at once. However, if the volume is very great, this may not be possible. Before the conversion is finished, a large backlog of transactions may build up and would be difficult to process.

2. Phased conversion, which builds up the data needed, a section at a time. This, of course, means that the data already gathered and placed on the EDP must be maintained through the use of the regular operating programs. The obvious advantage here is that the conversion workload can be spread over a period of time. The disadvantage is that you must run two jobs at once in the EDP, conversion and operations.

Each conversion problem will have to be considered in its own light to determine which is the best method to be used. A schedule should be set up for these operations and all personnel affected should be notified of the part they will play. It should be emphasized that a time cushion should be built into the conversion schedule. Experiences of past conversions have shown that, regardless of the planning done, unforeseen circumstances will arise that will slow the schedule down. If a proper cushion has been built into the schedule, these circumstances can be handled without delay to the operation as a whole.

At Mechanicsburg several months prior to installation of the "main frame", we arranged to receive the basic card-to-tape peripheral machines in order to commence an early conversion of our 34,000,000 punched cards. We recognized that

conversion of millions of cards to magnetic tape was a lengthy process and during the conversion period normal SPCC business operations had to continue.

To cope with this problem, we introduced parallel operations whereby card decks were converted to tape immediately following normal EAM operating schedules. The cards were then returned to the EAM room to continue our business operations. As daily changes were made to the original card files, duplicate transactions were reproduced and held for updating the tape records upon installation of the computer. This operation was employed throughout the preliminary stages of our "getting ready" for the computer.

CONTINUE CODING

Follow up on progress of coding to see that it is proceeding as scheduled. The Project Director should advise the Executive Committee of any deficiencies and suggest corrective measures to maintain the schedule.

ACCESSORIES, SUPPLIES AND EQUIPMENT - 7 MONTHS (CONTINUE CODING AND PROGRAM REVIEW)

ACCESSORIES, SUPPLIES AND FURNISHINGS

The Project Director should determine his requirements for miscellaneous equipment and consider the needs for:

1. Forms, internal and external
2. Card requirements
3. Magnetic Tape Requirements
4. Tape File Requirements
5. Furnishings, etc.

CONTINUE CODING

The Analysis and Programming Staff continue coding during this period.

PROGRAM REVIEW

During this month the Project Director and the Analysis and Programming Staff should initiate a review phase of the programming effort. The programming effort should be about 50% complete. If not, more effort should be channelled to programming. Pay particular attention to clerical errors, errors of interpretation of the block diagrams and completion of the End of File, error correction, and automatic Restart routine.

SITE PLANNING FOLLOW UP

Arrange for a meeting with the Public Works Engineers. See that reports of progress are made by all interested parties regarding letting of bids, selection of contractors, the ordering of necessary materials and supplies, delivery dates, and projected completion dates. It would be wise to review the floor plans in light of any changes made in components on order.

This is an area where a generous cushion should be added to be on the safe side. Having a room stand idle for a month or so is desirable rather than the other alternative - not having the room ready when the machine arrives. The odds are that at best the room will be ready at the last minute, in spite of your efforts to have it ready a little ahead of time.

MACHINE LOADING - 6 MONTHS
TEST DATA PREPARATION

MACHINE LOAD AND WORK SCHEDULING*

The Project Director should establish time schedules and priorities on data handling. He should schedule requirements for:

1. Main Frame
2. Auxiliary Operations
   - Card to Tape
   - Printer
   - Tape to Card
3. Pre-processing Schedules
4. Due-in and due-out requirements

*Hand out machine loading charts.

DESK CHECKING

Insure that all programs be given a thorough desk check. Experience has shown that about ninety percent of program errors are clerical and that the vast majority of such errors can be eliminated by proper desk-check procedures.

TEST DATA PREPARATION

Assist the Analysis and Programming Staff in the selection and creation of test data. Care should be taken to insure that the data used will test all conditions written in the program. The data should not be in large batches. This consumes machine time unnecessarily. Explain that a small quantity of carefully prepared test data will suffice.

COMPLETE PROGRAM REVIEW

The Analysis and Programming Staff should complete the program review phase during this period. In order to insure effective use of the program testing time, programs must be completely checked for elimination of the so-called "clerical" errors. A great deal of lost time during the program testing session is caused by machine stoppage due to program errors that should have been detected by a careful review of the program.

FINALIZE PLANS FOR CONVERSION - 5 MONTHS
REVIEW PROGRAM TESTING EFFICIENCY
REVIEW POWER AND AIR CONDITIONING

CONVERSION

During this month final plans for conversion of records to tape should be reviewed and approved by the Executive Committee. Operators and Programmers should all be apprised of their responsibilities for the conversion effort.

PROGRAM TESTING

The Project Director should maintain records of progress on program testing. He should note production statistics and compare them with the test center average. Keep track of types of errors encountered. Be especially watchful for errors in logic. Keep a current projection on the completion date of program testing. Insure that sufficient programs will be checked out so that productive work may start as soon as the machine is installed. Advise the manufacturer's representative immediately if additional time will be required, or if scheduled testing sessions will not be required.

REVIEW OF POWER AND AIR CONDITIONING

The Project Director should arrange a meeting with the Public Works Department Engineers to finalize plans on the machine room layout. Review delivery schedules of material on order. See that all persons involved are in agreement on completion dates for all phases of construction.

PROGRAM REVISIONS - 4-3 MONTHS
OPERATOR TRAINING
SITE CONSTRUCTION

OPERATOR TRAINING

The Project Director should arrange for training on data processing machine operations. Consider the following possibilities:

1. Operator training during testing sessions.
2. Console training at customer's office.

SITE CONSTRUCTION

Actual construction of the machine room should start this period. The Project Director and the Executive Committee should review pre-installation activities, and discuss progress to date and the work remaining. Review, with the Executive Committee, the plans for installation and the results to be accomplished within a specific period of time. Assure yourself that all members of the Executive Co mmittee and top management are in agreement with facts presented and objectives to be reached.

FOLLOW-UP ACCESSORIES

The Project Director should follow-up on the delivery schedules of supplies, accessories, and furnishings. Review delivery schedules on tapes, punched-card equipment tie-ins.

Tape Cabinets
Reels of Tape
Ribbons
New Forms

FINAL PROGRAM TESTING - 2 MONTHS
PROGRAM REVISIONS - CUTOFF
PERSONNEL ASSIGNMENTS
PHYSICAL REQUIREMENTS

ASSIGNMENT

The Project Director should assign personnel to specific duties and responsibilities. He should insure that all persons involved are aware of details of their jobs. Discuss shift make-up if required. Review plans for:

1. Operation of the installation.
2. Future work loads.

See that the strength of the Analysis and Programming Staff is spread over these areas.

PROGRAM TESTING

Plan to complete check-out of all programs necessary for use on installation date. By the end of this month all details on machine procedures should be completed, tested, and ready to run.

PROGRAM REVISION

Since the program testing sessions are due for completion during this month, the program revision phase should be completed. It may be necessary to "put a freeze on" changes to programs until after the machine is installed and operating.

PHYSICAL REQUIREMENTS

Insure that you are able to install the machine when it arrives. Arrange for testing of air conditioning equipment two weeks prior to machine arrival. Final balancing of the system may be done after customer engineering has turned the power on.

FINAL DETAILS - 1 MONTH

REVIEW OF FINAL DETAILS

The Project Director should insure that the following points have been completed:

1. Power installed and tested.
2. Air conditioning installed and tested.
3. Control devices (humidity and air conditioning) installed and tested.
4. Insure delivery of necessary tape reels.
5. Follow-up shipping details with IBM Traffic Department regarding arrival date of machine.
6. Follow-up on arrival of customer engineers.

About four months prior to installation of your machine, it would be wise to publish an "EDP Status Report" for your management and supervisory personnel. About this time everyone is getting anxious, asking a lot of questions. Everyone wants to know when their application will go on the machine. You can gain a great deal by doing this. I have a copy of our status report. The first part of the book contains a summary of the EDP status which could be easily understood by all supervisory personnel. We made about 100 copies of the summary but a very limited number of copies of the whole book.

Ladies and gentlemen, it's been a real pleasure to be here today. I trust that the information we discussed this morning will be helpful to you in the event you should install an EDP machine. Your pre-installation planning chart can be used as a check off list of things which need to be done and which need constant attention. I want to repeat again, that you should determine what your objectives are as soon as you can and continually strive to attain these objectives. A computer has to be of paramount interest to top management--and management must start with a definite point of view on the final results it wants to accomplish. It's not too difficult to install a successful computer operation once worthwhile objectives have been determined, provided the people responsible for the installation are not afraid to make decisions, to sweat a little bit and to never lose sight of the objectives.

I leave you with this thought; you may buy the most expensive machine on the market, but it will be no better than the people behind it.

GRADUATE SCHOOL
U. S. DEPARTMENT OF AGRICULTURE
SEMINAR ON AUTOMATIC DATA PROCESSING
FOR FEDERAL EXECUTIVES

Ecology and Techniques of Operation

By Frank W. Reilly
Special Assistant
Assistant Administrator for
Management Services
Veterans Administration
Washington, D. C.

I ECOLOGY OF ADP

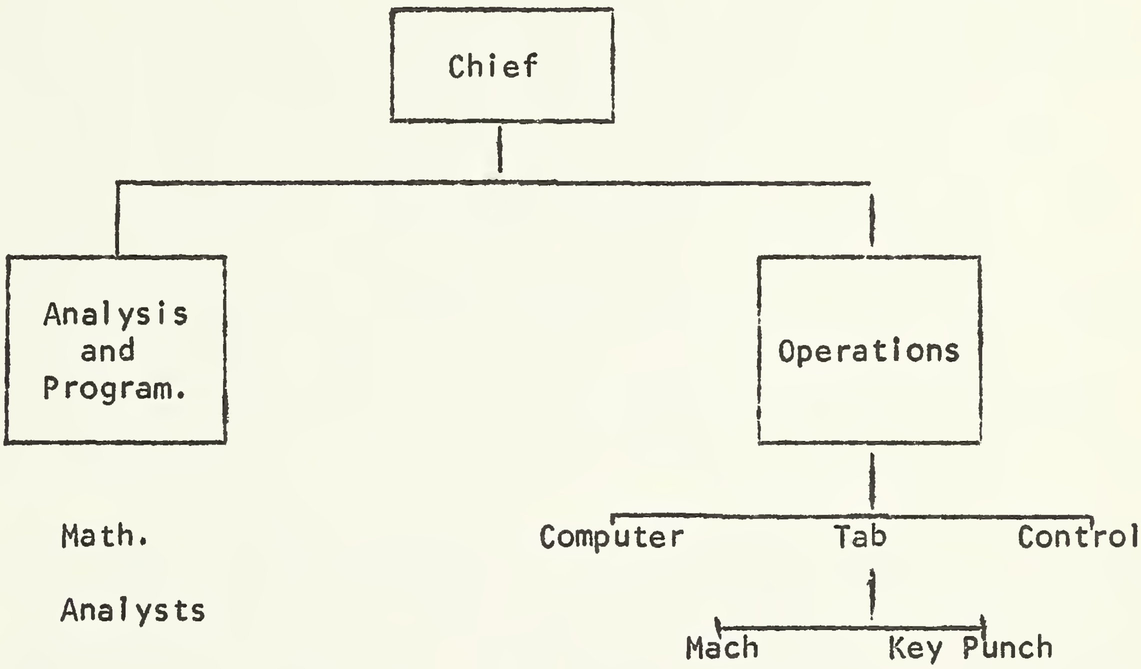
When people discuss the functioning of a data processing center, they frequently have a mental picture of a large room filled with blinking and buzzing machines. They tend to forget that this equipment is used by human beings to provide requested services to a much larger group of human beings - their customers. To understand adequately how electronic digital computers can be used successfully we must appreciate the necessity of having a fructuous environment in which both these groups utilize the ADP technology.

The working atmosphere which provides fruitful ADP should be found in those data processing centers that are examples of productive endeavor. In my candid opinion few of the hundred of extant systems could be described as living advertisements of the best in ADP. Most of them are going through what Canning (1957) calls the two-step approach and what I euphemistically call the two-step tango. That is, you get a computer with the realization you know nothing and say: "We know that at first our computer is going to be greatly wasted, but we are taking the envolutionary approach. It's at the second step, after we've been educated that we are going to start making real use of the computer." I am afraid most management has chosen this way, and a goodly number are still in the first step. But some organizations decided at the time of the feasibility study that they didn't want to two-step tango, they wanted to one-step. They said: "We are going to make a success of this thing from the beginning." It's from an analysis of these data processing centers that I have gleaned the following two key concepts:

First, they have a working philosophy that evolves around the concept that operational integration and not hardware concentration is more important. That is, they devote their major concern to systems performance. (As an aside may I say that this integration problem is endemic to all phases of 20th century living. The daily paper headlines it as a social problem, It's one that strikes at the heart of the whole military organization in this country. Also we face it in the joint conduct of military and foreign affairs, not to mention conflicts between the military and space people. It is a problem that we face in our cities. How we integrate the varying demands and needs of local, national, international and interplanetary societies is high on the list of priority projects for our leaders to solve. So this philosophy of operational integration is not peculiar to the needs of ADP.) The second key I believe is a psychology of service. The people in an ADP installation must take the attitude that service is their most important product. I do not wish to convey the impression that I mean service as a mere mechanical handout from a machine. The production by a computer of so many hundred pages of such and such report may well be a disservice. Service in my context means the computer user must be aided from the time he initiates the project until it is finally completed. These two concepts can become meaningful as they are related to the organization and operation of the DPC.

Below there is an organizational chart for a DPC. Its chief has two major organizational units paralleling his main functions, analysis and programming and operations. I have not broken analysis and programming into a number of separate functions. You probably have seen diagrams that break these down into various sub-groupings. From past experience I prefer not to do this and viewing future trends it seems unlikely this will be desirable. These are my reasons.

DATA PROCESSING CENTER



Let us say we have a DPC that handles both scientific and business type applications which means it is staffed by mathematicians, statisticians, and analysts. As I see it, when working on a given problem, a mathematician, statistician or analyst will take it all the way from beginning to end. He will do the analysis, the flow charting, he will do what is referred to as coding, the debugging on the machine and finally he will do the preparation of the instructions for the computer console operator. If you question the validity of this approach today, you might agree that future developments of machine independent common languages such as COBOL and ALGOL will eliminate all this brute force machine language coding - using these language systems and automatic compilers, the man who analyzes the problem can perform the systems study, make his flow charts, and be able to write their logical propositions in fairly simple English or psuedo-English which the machine can accept. Therefore the need for the tiers of specialists such a analyst, programmer and coder now seen on some organization charts will be eliminated. These three tiers will converge into one function as we become proficient in using common languages.

I have preferred this approach in the past because it enables management to fix accountability and get more efficiency out of a man by giving him full systems responsibility from start to finish. In this manner he can generally obtain a close working relationship with the project sponsor and thereby provide personal service. What I'm saying is that I saw more advantages in a generalization of labor than I did in the specialization of labor. Many people would disagree for what they consider legitimate reasons, but my choice stresses the twin principles of integration and service.

One of the important axioms of DPC operations is that you do not create hierarchies of people. On the operations side of our chart, we have both computer operations and the regular tabulating or EAM operation, which is subdivided into machines and key punch. There is also control and clerical operations. Here at the bottom of the chart is the lowely key punch operator and here at the top is the mighty mathematician or analyst. There is latent personal friction and animosity in this structure since the pedestrian can resent the status of the equestrian. These resentments must never develop if we wish to avoid having this costly equipment become just an expense without providing the expected benefits. Nowhere is it truer that a chain is no better than its weakest link then in ADP. The entire staff has to be made to realize that every employee has important duties, that without a keypuncher and good input everything else becomes worthless; without adequate data control before going on the computer you are just trying to purify garbage and you cannot verify the validity of the output (and this is not uncommon); that your EAM group are a necessary part of both input and output control plus providing ancilliary services which many users consider indispensable. Finally we come to the actual operations of the computer and here sits the console operator, one of the most important people in the entire DPC. Sad to say the job classification standards established for this position were inadequate three years ago and they have become more out of line since. The net result is that frequently your best programmer becomes a computer operator, because he is the only one that really has the savvy and the know how about computer operations and programming to get the computer to do what the programs intend.

All of the personnel become important in an operationally integrated DPC, if any degree of efficiency is to be achieved. When you also consider the reason for the existence of the DPC, service to the user, this integration achieves further significance. If everybody in this shop isn't pulling together how can management ever meet its commitments to the users. There are many various specialities in a DPC which must be brought to bear on a particular job. If anyone of them fail to perform properly the end product will be delivered late and its quality questionable. We have a serial work situation here in which everybody has to feel that they are part of one team producing a single output at a definite time. In order to convey to the individual employee where he fits into the overall scheme of things, management must explain the utility of the projects handled by the DPC and the importance of what the employee is doing. I can illustrate this point by an experience of my own, while in charge of a DPC that was providing computing services to earth scientists. It happened that one of our major jobs was in conjunction with an X-ray crystallography laboratory. The geochemists had a sizeable workload of keypunching which required a high degree of accuracy since the X-ray analysis of atomic structure depends upon intensity measurements recorded at lattice intersections. Since this work was important and of a continuing nature I thought it would be worthwhile if the key punch operators had some concept of what the data was that they were punching. I gave them a brief non-technical description of what crystallography was, the reason for the digital computations and the great value placed upon speedy and accurate key punching. I was led to believe that as human beings they appreciated someone taking the effort to explain crystallography, how it fitted into the overall functions of their agency and what its contribution was to an improved society. At the same time they better realized just how important it was for them to key punch in a superlative manner. The crystallographers and other scientists often remarked about the unusually high degree of accuracy they obtained from our key punch room.

There is one final sequel to this illustration. During the years that I had this DPC our total turnover of personnel was one mathematician. This was during a period when most DPCs hardly became acquainted with their personnel before--they

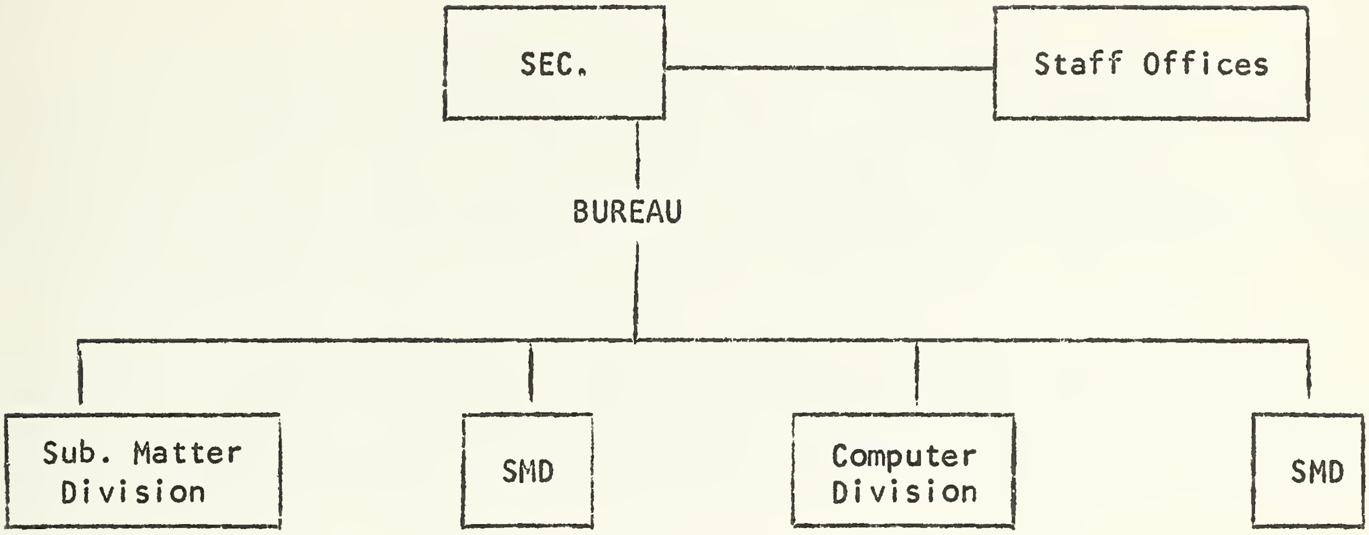
moved on to greener pastures, generally in industry. I have come to believe that people are looking for something more than just a bigger paycheck every two weeks. They want to feel they are an integral part of a team, (that hackneyed and badly abused work) which is making a worthwhile contribution to this society. They are not living just to make money so they can buy a fancier car or a larger house or a sharper suit of clothes. Instead most people basically are looking for job satisfaction and unfortunately few of us supervisors are providing it. Forget the do-good approach and look at it strictly from the standpoint of good-doers. You want employees who produce at the maximum and to do this they must know what the job is all about and where they fit in. Since automatic data processing equipment has become the synonym in the popular press for a mechanistic way of life, it behooves all who are involved in this business to prove the converse is true.

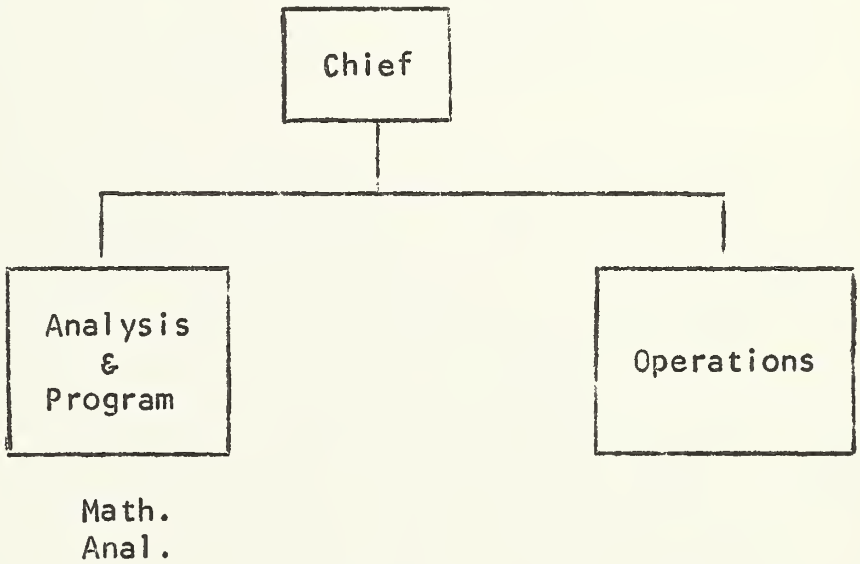
The first organization chart displayed only the internal relationships of a DPC. The chart on the opposing page relates the DPC to other parts of the parent organization. The actual relationships to be discussed are those between the subject matter divisions and the computer division (DPC). The latter group exists for the purpose of serving the former, the people and the problems. The question arises as to how deeply subject matter personnel should become involved in the technology of data processing, particularly programming. In the case of subject matter people who are going to have a long continuing computation project, I think that they should learn programming. They can still call upon the DPC programming staff for assistance, say for improved techniques in numerical analysis. But they should at least have a basic knowledge of programming for it will be a definite aid in enabling them to better organize their own work. Now for those subject matter people that have smaller projects or one-time projects or those which occur only twice or spasmodically. The best approach is to have problem formulation at the subject matter division and then call upon the DPC for programming. In other words, have the DPC programmers (mathematicians and analysts) work with the professional subject matter staff.

I should probably say a word about the magnitude of the programming staff. For an average size governmental bureau, found in Agriculture, Interior, etc., a staff of four or five mathematicians and a like number of analysts is generally adequate because you want to develop within your subject matter divisions, the people with the knowledge and the interest to really carry this program over. The DPC staff, if they have the proper systems orientation, should do much to educate everybody. Further, if they understand the service concept they will really feel that it is their function to be doing this type of training. So I think the relationship between the DPC and the subject matter divisions should be a two-way street with the former aiding the latter where possible, either in actually doing the programming or giving technical advice and the subject matter people with the problems bringing them to the DPC.

I have ignored the relationships between the staff offices, at the top of our chart, with those of the bureaus and their divisions, both subject matter and computing. It is difficult to describe these relationships because for the most part they have been deficient and in some agencies non-existant. Nonetheless, it is generally recognized that corporate SDP activities of the large industrial concerns, Ford, Chrysler, Westinghouse protend the future for these staffs. As a minimum the staff offices should perform the two functions of information exchange and agency forward planning. In addition, they are expected to pay an operational role in ADP.

AGENCY RELATIONSHIPS





The key to the role of the staff offices is the fact that ADP cuts across organizational lines and in the process revamps processes, organizations and concepts of management. A top staff office is expected to be free of the program bias position to see what has been successful and unsuccessful in the way of computing applications and actual DPC performance. Equipped with a free mental outlook and a host of information this office should do all in its power to assist and encourage those parts of the organization who have or conceive they have possible data processing applications. It should be in the forefront of advanced management planning for not only future ADP applications but also for the effects these applications will have upon managing the substantive function of the agency. We could hypothetically say that if ADP has the promise of greatly assisting animal husbandry then the staff office in the Agriculture Department should be studying the potential of this upon the bureau concerned. Most Departments will eventually feel the effects of ADP in their prime functions.

The operational role of staff offices in ADP comes about because of the growing power and capacity of equipment as well as the disappearing significance of certain organizational lines and entities. For example, the corporate DPC of Chrysler

Motors can serve all of the major automative divisions. Likewise, it does not take much imagination to see that a machine with the capacities of the Larc or Stretch could serve more than one bureau in a Department where problems lack the computing complexity of neutron flux patterns (the prime application for both machines).

For staff offices to operate in a fashion which would enable them to perform the functions I suggested above means they have to have excellent and knowledgeful relationships with all parts of their organization. From the DPCs they can learn the vageries of operational problems, from the subject matter divisions they can learn of present and future applications. They should be in an excellent position to carry the gospel of ADP from and for all of data processing management throughout the organization.

Implicit in the concept of a DPC being service oriented and always striving to operate on an integrated basis are the requirements for (1) special qualities in key personnel and (2) the necessity for these personnel to be "systems design" conscious in their analysis of sponsor's problems. A resume of each point will import by meaning.

The key personnel in the DPC are the manager, the chief of analysis, and programming, the chief of operations, the senior mathematician and analyst, and the senior computer operator. These people must work together as a unit, obviously, but in addition each should bring to his job certain qualities that can enhance the success of the organization.

The manager of the DPC must have a balance between technical data processing competence, systems engineering and management ability, with the latter two being weighed most heavily. He will, understand the computer technology including hardware specifications, programming and general machine applications. This man needs a good understanding of what makes people tick, both his own employees and his clients. He must be able to deal with all the subject matter divisions even though he may not be intimately familiar with their potential applications. He should be capable of grasping the potential of their ADP applications in a manner which convinces them of his sincerity to help and the ability of his staff to deliver end results. Above all this man must be completely above board in his dealings with the clients. If things go wrong, and they inevitably will, this man must not look for a scapegoat to pin blame upon. There is nothing that can kill effective use of ADP quicker than a DPC manager with a personality that seeks to build himself at the expense of others either by grabbing glory or fixing blame. And finally this individual should be planning for the future at least three years hence. Once the DPC has been firmly established this last activity should become his primary job for today's success is built upon yesterday's plans.

The chief of analysis and programming is the key technical man in the organization. He must possess a fund of knowledge on systems analysis or mathematics, be intimately familiar with programming, and be capable of dealing effectively with the many customers of the DPC. Last and not least, he must supervise his staff in the manner to encourage creatively while yet meeting work commitments. There is nothing more elusive than estimating analysis and programming time on a job, but this man must somehow he reasonably successful in so doing. Generally, a wealth of past experience is the only manner in which he is able to base his estimates.

Our operations chief should not be a person who is allergic to aspirins, since an unending supply of headaches is his lot. Here are a few of his musts: he must insure that all information going into the computer had adequate controls; he must meet schedules internally and externally; he must keep his own people happy and not out looking for a new job; he must realize that certain projects have to be given priority on the basis of both need and importance. Many of these musts are true for any supervisory job. The function is more critical here because of the large amount of money involved and becuase of the unbelievably complex work relationships with other offices throughout the organization.

The senior mathematician or analyst is expected to have an excellent technical background as a minimum. The quality most looked for here is a history of successful work and personality that thrives on effective accomplishment. A good technical man who has no real interest or drive in getting a job finished does not belong in this position.

The senior computer operator, or console operator as he is more commonly called, is one of the most underrated jobs in automatic data processing. Especially during the first year of operation, the skills of this man can make a difference between success or the loss of literally hundreds of good hours of computer time. Realize that much of your input data will be inaccurate initially, that the maintenance of the digital computing system seldom stablizes before six months, that programmers will require much help in debugging their programs, that debugged programs will continue to develop "hangs" for many months and that your entire DPC has to learn to work as unit before it can really produce effectively. During all these travails, a top notch computer operator in a large system is easily worth his weight in gold (I calculate this at 150 lbs X 16 ounces X \$32 = \$76,800) for all the wasted effort in machine time and manpower he will save.

Now that we have employed a staff we must be concerned that they always remember to think in terms of "system design" rather than "hardware orientation" when dealing with their clients, the subject matter divisions. This means that first and foremost they must conceive that a prime part of their "service" orientation is to assist and encourage the development of improved management systems and the optimization of the agencies function (substantive applications of both business and scientific type applications). This does not mean to imply they should view the problems presented to them just in terms of getting applications on the computer. Rather they should look to the extent possible to see if the project makes sense from a systems standpoint. A subject matter division retains basic responsibility for the performance of its activities, but it will generally welcome an honest attempt to help it do a better job. Thus, I do not believe it is improper for a DPC to question the total problem formulation of a sponsoring sub ect matter group. An illustration will demonstrate my point.

A project sponsor comes to a DPC and asks for assistance in analyzing a potential ADP application. His problem involves taking input data obtained from semi-quantitiative chemical analysis, subjecting it to statistical mainpulation which would produce correlation coefficients who significance could then be determined. After a great deal of programming effort and a very sizeable expenditure of computer time the DPC produces table of statistical correlations. Upon examination it is readily observable that only a minuscule number of correlations exceeded .3 and few of these approached unity. Of those correlations exceeding .3 there is no discernible statistical relationship. In other words, thousands of dollars have been spent on producing irrelevant statistical information. From this the DPC learns that in the future it will have to raise some basic questions with the project sponsor about the following:

(1) Nature of his experimental design
(2) The reliability, validity and accuracy of his input
(3) The extent to which he has sample tested his hypothesis before introducing large amount of data for analysis, and
(4) The kind of criteria check or decision making he can build into his computer programs

DPC's cannot affort to ignore their responsibility in seeing to it that project sponsors get the maximum benefit from money spent with them. True they cannot unilaterally refuse to do what a client requests, but they can point out the minmal "systems" design to be expected of a well run project. If this assistance is given in the proper spirit the subject matter divisions can be expected to respond in the manner desired.

The only remaining matter of significance that effects the immediate relationships between a DPC and its clients is the problem of language barriers. First lets us look at it from the side of the project sponsor.

He has spent his professional career in learning a language which is meant to provide a mental shorthand in explaining complicated phenomena. This is true to both scientific and business type applications. He wonders if the DPC personnel can understand his problem enough to help him. The answer is that a good mathematician is well founded in basic scientific concepts and a good management analyst is similarly founded in basic management principles. Both require the project sponsor to formulate the problem in sufficient detail for them to adequately perform an intelligent analysis. This does not mean that they will have to become experts in a particular field of technical endeavor for they will rely upon the sponsor for elucidation of points of technical detail. Where a project is of a continuing nature it is most desirable if someone on the staff of the sponsor becomes familiar with programming so they can realize the extent to which the computer can assist them in their work.

The other side of the language barrier concerns the sponsor understanding what it is the DPC expects from him and the techniques they are using to solve his problem. The sponsor should be given a check list outline of what the DPC expects from him in the way of:

(1) Input specification
(2) Problem definition
(3) Problem consultation
(4) Problem test
(5) Output specifications.

The programming staff should then do its utmost to explain to the sponsor their approach in solving his problem, including an explanation of their block diagrams, if he is interested. If the programmer is coding his instructions in machine language it is not practical for him to go into explicit detail. In the future, as cryto-English and pseudo-English common machine independent languages become popular, the problem sponsors may eventually even become familiar with the coding process.

In concluding this section, I hope I have gotten across the point that a DPC is not a place where machines congregate. That instead it is peopled by human beings who desire to provide service to the entire organization and whose purpose is to insure maximum utilization of the ADP concept by promotion of the principle of systems integration. In addition they strive to achieve efficiency through operational integration in their daily data processing function.

II PRACTICALITIES OF ADP OPERATION

The most practical question regarding data processing operations asks, "How do you get the best use out of your ADP system?" In more specific terms this question has two parts. First, how do you obtain the most efficient utilization of ADP equipment? Second, how do you enable those who need data processing services to make use of them? Assuming that potential project sponsors are desirous of using ADP services to the utmost, the answers to our last two questions hinge upon certain techniques of operating management. Key to the implementation of such techniques are the establishment of adequate budgetary policies and cost reporting systems. A good budgetary system conbimed with rational cost accounting practices can encourage the best use of a computer while providing an incentive for efficiency.

In my opinion, a DPC should operate financially on a reimburseable basis. This is particularly advantageous where it functions as a service center type of activity. This means that each project sponsor (computer user) would be charged directly for all services received except where such charges are cumbersome to maintain or where they discourage the best utilization of the system. A further expalnation is needed to clarify these statements.

The purpose of a DPC is to provide service, to do a job better and to perform work that was not possible before the advent of ADP (e.g. three dimensional neutron flux or safely maintain minimal large scale inventories). Thus the budget people must understand, in a general way, the utility of the computer applications and the directional trend for this kind of work. A reimbursable budget for the DPC requires that each major organizational unit requiring data processing services designate certain funds for this purpose. This device has a built in safeguard in that each user has the monetary incentive to (1) put on the computer only those problems he believes are worthwhile and (2) be actively interested in complying with operational efficiencies developed by DPC. Simultaneously this reimbursable budget becomes a useful device for forward planning. The manager of the DPC can call together all the users to determine their future workload requirements which, of course, determines staffing patterns, equipment additions or changes, etc. As a by-product, it also has the advantage of forcing the subject matter personnel to give serious thought to their planning for ADP applications. A financial plan with its twins, source of funds and application of funds, cuases many echelons of an organization to think in concrete terms about long range projects.

The essence of financial planning is knowledge of current costs and it is in this context that cost reporting assumes its role. There are about three major expense classifications chargeable to a computer user.

First there is computer hours. This should be a directly chargeable, single, flat cost per hour, which includes peripheral equipment, operators cost, and the overhead expense of clerical and supervisory employees as well as supplies. This direct charge could run as low as \$50 per hour for smaller systems to \$600 per hour for the large ones.

Second, the tabulating equipment (EAM), which is not associated with the computer. These machines are often used to support computer installations by providing a ready access to pinters, collators and the like. This equipment should be directly charged to the user on a flat cost per hour per type of machine. The flat hourly cost would include equipment use, operational labor and an overhead charge covering supervision and supplies.

Third item is analysis and programming services. These are the expensive man-hours which convert the sponsor's problem to a formula and language the machines can accept. Though all the other items of cost have been direct charges to the user I believe this item should be paid out of a general fund. My reasons for this are several. One point is that it is most difficult to accurately estimate programming costs a year in advance as budget planning requires (surprisingly enough computer hours are much easier to estimate from just a feel of the job). Another reason for this approach is the need to provide sustenance for nascent applications. You must encourage project sponsors to broaden their concepts of how ADP can help them. It seems invariably that they cannot afford to pay for programming since many subject matter people do not have applications that can show immediate tangible savings. There is no better way to eliminate their participation then to charge programming directly to them. If you will recall, the opening paragraph of this section asked, "How do you enable those who need data processing services to make use of them?" I think here is where budget people must have the foresight to explain to management why expenses for mathematicians, statisticians, analysts and programmers should come from the general fund. If the DPC is established to help solve the substantive problems of an agency the usual budgetary policies cannot be allowed to starve out potentially good computer applications. We must avoid the mistakes of the banking fraternity who formerly would only lend money to those wealthy enough not to require a loan.

On this last point a comment is worth noting. Everyone must realize that all applications will not receive a payoff either in the short or long term. There may be some horrible failures, for man is fallible. It is only by evaluating the overall program that you can appraise the success of an ADP program. The failures must be compared to attainments in determining if the operation is successful and not just pick out either one or the other as your yardstick.

Our cost reporting system serves as a yardstick of performance. One of the operating details of a DPC, which will be reflected immediately through the costs reports, is the skill with which work is scheduled. When we remember that computer costs range from $50 to $600 per hour this skill can be ranked as being indispensable. There are a few essential matters to be considered in regards to its exercise. First it must be recognized that the work begun today by the programming. This fact requires that priorities be established for each job. Criteria determining the ranking of a project involve (1) the nature of the work, (2) the estimated computer time and its availability (3) available programming skills (4) extent of debugging and availability of computer time (5) available project funding for computer time after debugging (6) ease with which a project can be slotted between production jobs (one time vs. regularly scheduled) to maintain just a few of its most obvious ones. Once the programming practices have been established work can be started to plan for computer time.

A degree of flexibility in scheduling projects for the computer is quite helpful. The constitution of a DPC's workload has a major effect upon this flexibility. A computing installation which processes only administrative applications on periodic fixed dates has much less flexibility of operations than a DPC which handles both regularly scheduled and intermittment type projects. Or to phrase the situation in another way; those installations which have a mix between both scientific-engineering and administrative applications are in a position to operate with greater efficiency those having only the latter. This is because of the four types of workload, regularly scheduled jobs, debugging, one time jobs and continuously but intermittenly scheduled jobs, administrative applications generally fall into the first two categories only. Thus once a DPC has completed the bulk of its debugging, all that remains are an unending succession of fixed deadline projects. This leaves no squeeze room, or margin for machine

failures, operator error, programming hangs, job reruns and the like, which inevitably occur in the best run DPC's. Thus the management of the DPC must set aside a certain number of hours per day or week for contengencies. If they fail to materialize he is hard put to utilize the extra time. On the other hand he dare not make omit such a provision if it is mandatory that all deadlines be met.

This condition militates for a better distribution of the types of workload in every DPC. At the very least it is an encouragement for agencies to look to their substantive program to see what applications are lying about for ADP to aid. This may sound like the back door approach to promoting the rational use of ADP in an agency's main program but I know of a case where it has succeeded handsomely. This private concern had a DPC devoted exclusively to administrative applications. The center management realized that (1) he was always under the gun schedule-wise and (2) that he was not furthering the best intents of his company by ignoring all of its main problems. Over a two year period the manager slowly changed workload emphasis until he finally achieved a 50-50 balance between administrative and substantive applications.

There are a few techniques of actual daily operations that I might comment on before concluding. Always try to schedule ADP work at least a week in adva ce. Consult with the individual users about what their requirements are. Check with the head of programming for debugging time, contact the engineers for possible extra maintenance or field modifications. Finally have at least two projects in a standby status so you can use them or not as the situation demands. In the practice of actually putting the job on the computer make cettain that all card and tape files are accessible, that complete operating instructions are provided the computer operator and that a computer log is religiously maintained. This last requirement can save more blood, sweat and tears than twenty speeches by Winston Churchill. A good computer log tells every body all pertinent facts about what is happening in the computer room. If well kept, its data is invaluable for analysis of past operations which provides the basis for future planning.

GRADUATE SCHOOL
U. S. DEPARTMENT OF AGRICULTURE
SEMINAR ON ELECTRONIC DATA PROCESSING
FOR FEDERAL EXECUTIVES

Relationships With Equipment Manufacturers

By F. P. DiBlasi, Jr.

I am sure there will be no disagreement with a statement that developing ADP systems is an expensive business. Moreover, I am sure everyone will agree with the statement that it is a difficult business. Therefore, let me use these truisms to support the premise that the high cost and complexities of ADP demand the concerted and cooperative efforts of all groups of people who are concerned with the planning, selection, installation and operation of an ADP system. This thought can be expressed in one word - TEAMWORK. It implies each group contributing to the common objective, in the area in which it is equipped best to contribute, to the extent that the contribution is needed and not more. Of course, there may be several different groups involved in developing a given system, but one group which always is involved is the equipment manufacturer. The nature of the role he is to play, the extent to which he is to be involved, his responsibilities as a teammember will vary from system to system and rightly so, depending on the planned approach that management takes. The point remains, however, that there is a vital relationship between the customer and the vendor; a relationship which must be understood by both parties and accepted by both parties in good faith in order to assure a winning team effort.

There can be no doubt that the customer has the responsibility to design and install an ADP system which will attain its goal. He cannot delegate this responsibility, but he can and should enlist the assistance of the best qualified sources of ADP knowledge and use them on the team in positions where they can serve him best. However, remembering his responsibility, the customer will always be the team's manager and call the signals. So, the equipment manufacturer's job is to help the customer. The manufacturer welcomes the role of helper because he knows: (1) that he cannot do the whole job, (2) that he can make important contributions, and (3) that his contributions will help assure a successful installation and a happy customer, which is the principal ingredient of a recipe for more sales--his ultimate and honorable motive.

The relationships with equipment manufacturers will not be the same throughout the lengthy time required in developing, installing and operating an ADP system. Therefore, we have segmented these relationships chronologically into four periods:

1. Before Equipment Selection
2. Equipment Selection
3. Before Installation
4. After Installation

1. Before Equipment Selection

This period is characterized by the important functions of Organizing, Educating and Planning. It is a dangerous period in that these three vital functions sometimes are passed over lightly either because of lack of appreciation of their importance, or because of an almost passionate desire to

possess this most glamorous amazon of automation--the computer. This is the unglamorous period of developing and selling precepts for ADP, of recruiting and selecting people, of schooling staff members, of visiting and evaluating, of soliciting advice and assistance, of indoctrinating management officials, of studying and designing concepts of better information systems, of documenting, flow charting, justifying and making numerous decisions in the face of as many opinions. At best, it is an extremely difficult period, yet a tremendously important period in which the foundation for the ADP endeavor is built.

The equipment manufacturers can be helpful at this time if they are used to perform services which are necessary to organizing, educating and planning, and which they are qualified to perform best. Generally, manufacturers have a wealth of experience to draw on; probably more than any customer can compile. It should be recognized, however, that much of their experience has been derived from observing mistakes. This is valuable experience but often is manifested in what not to do rather than in positive terms of what to do. The ADP customer will be wise to listen carefully and decide carefully just what he should request manufacturers to do for him. ADP is a competitive industry and most vendors are prepared (perhaps reluctantly but of necessity) to take as active and as complete a part as the customer demands. In this first of four periods, the customer should condition his own organization to carry the ball. For instance, proper indoctrination of top management is a critical problem and must be composed of a delicately balanced mixture of equipment information and the true concept of cost and timing to obtain a full realization of top management's responsibilities for support and patience. It is believed that this purpose can be achieved best by highly respected, in-service personnel who will be able to address themselves to management's most pressing problems in terms of management's specific concerns. The equipment manufacturer who is placed in the role of indoctrinating top management is apt to overemphasize equipment, which is his forte, and to neglect management's principal interests with which he cannot be expected to be thoroughly conversant.

Another important function of this first period is selecting employees for the ADP program. In this connection, some manufacturers have developed tests which are available as aids in discovering aptitudes for different types of ADP jobs, especially programmers. Such tests have been used widely; often with excellent results. A close working relationship with the manufacturer in testing aptitudes should be a valuable service to the customer.

Usually, three types of schooling are required during the first period--systems analysis, programming, and equipment appreciation. This is an area where the manufacturers and the customer should work together very closely. In programming and equipment appreciation, the manufacturers are unsurpassed. Often, the manufacturers, if requested, will evaluate the ability and progress of the students and report this information to the customer for his use in placing people in his program. Some manufacturers have developed systems analysis courses for those customers who have no means for conducting their own. The manufacturers agree, however, that systems analysis training can be conducted best by the customer, inasmuch as the courses can be tailored to his own particular needs. In this connection, it should be recognized that manufacturers are equipment sales organizations, not management engineers.

We have touched on the functions of Organizing and Educating with respect to relationships with manufacturers. Let's turn now to a consideration of those relationships in the planning of data processing systems prior to selecting equipment.

The familiar term is "Feasibility Study," which has come to mean almost any type of ADP study depending entirely on to whom one is speaking. The Navy briefly defines Feasibility Study as "an over-all reappraisal of the management information and data processing system requirements of an organization. Its objective is a determination of the extent of the probability that operating economies or improved management effectiveness can be realized through the adoption of system modifications to meet management needs." Usually such a study can be done within 60-90 days. So we are talking about a rather short term study to find out if we are warranted in looking at more detail with respect to specific applications for ADP and developing a sound basis for a valid selection of equipment. Navy calls the second step, the Applications Plan. But first, let's examine relationships with manufacturers during the Feasibility Study.

Since the Feasibility Study does not deal directly with the question of equipment, except to conclude that probably some form of mechanization will be (or will not be) of benefit, there is little need for coordination with manufacturers. The customer is not yet in the equipment stage of ADP systems planning, It is true, however, that the feasibility study group should include people with knowledge of mechanization, but even they are concerned primarily with management needs for information on a broad basis. Some organizations have taken a completely opposite approach by inviting outside concerns such as management consultants or equipment manufacturers to make the Feasibility Study. Such an approach is less common now than it was a few years ago when ADP was brand new. I cannot say that this approach has ever been successful; but, I can say that in some cases horrible results were obtained. It seems obvious, I believe, that the Feasibility Study can be conducted better by in-service personnel. They best know the structure, the processes, the mission, and peculiarities of their organization. OUtsiders must either work without that precious knowledge or spend a great amount of time trying to acquire it. And, of course, it has become a trite but true saying that when the outsider leaves, all knowledge goes out the door with him. Sooner or later the ADP customer must build his own capability. If he does not have it now, he would be well advised to postpone the Feasibility Study until he does have that ability.

The second step in system planning is the Application Plan, which is a more thorough study undertaken after the feasibility study has indicated an affirmative finding. Highly competent ADP technicians are required for this work which is directed toward the utility of equipment and optimizing the flow of management information. It involves the study of system inputs and outputs; examining frequency, volume, format, necessity, etc. It produces narrative descriptions and flow charts of proposed systems in sufficient detail to facilitate a determination of the general range of equipment which will be required. From this brief description of the Applications Plan, it may be inferred correctly that equipment considerations play a greater role than in the Feasibility Study. It should provide all the equipment information necessary to embark on a procedure for evaluating

and selecting the specific machines which will be best for the job. It follows, therefore, that the study group will be in more frequent contact with equipment manufacturers to determine the specific capacities, speeds and other capabilities of the various makes and models of ADP equipment to insure that the system being developed is amenable to ADP methods and to have at least a general idea of what implementation will cost. Still, the task is a systems matter and should be done by in-service personnel rather than by outsiders. At the conclusion of the development of the Applications Plan, the ADP manager will have a set of system specifications which can be used for the purpose of selecting the most appropriate equipment.

2. Equipment Selection

Perhaps the basic principle of equipment selection is equal opportunity to all manufacturers. Regardless of whether evaluations are made privately by in-service personnel or as a result of proposals solicited from manufacturers, it is in the best interests of the user to consider all manufacturers and to afford each the same chance to furnish the equipment for the system. In short, equal opportunity will foster competition among manufacturers, which in turn will result in lower prices, better terms, and most important--technological progress. Much progress has been made already, but much more must be made before ADP equipment can facilitate true integrated management information systems. Hence, the need for competition.

The usual practice for selecting equipment is to giveeach qualified manufacturer the same specifications and system requirements, requesting him to propose equipment to do the job. We say "each qualified manufacturer" realizing that the nature of the system requirements themselves, or a necessary installation date, or the non-existence of a planned machine, could be sufficient reason to regard an ADP manufacturer as unqualified. Except for such reasons, all manufacturers should be considered even though one may not choose to submit a proposal.

Proposals by manufacturers should show in detail the methods by which they propose to accomplish the workload, time estimates and costs including one-time costs. Because these estimates are so important, manufacturers should be given ample time to study the system and prepare their proposals. Any privilege extended to one manufacturer should be extended to all.

The customer should validate carefully all time and cost estimates submitted to preclude the possibility of selecting equipment which appears to be best only as a result of erroneous calculations, or as a result of calculations which are based on other than actual system specifications. This period of evaluating and selecting equipment is one when liaison between the customer and the manufacturers will be fairly constant. This is when the manufacturer becomes a part of the team, and when he should be encouraged to suggest any changes which will improve the design or operation of the proposed system. His rather extensive experience, if properly brought into play, should be very helpful to the customer. If we recognize that the manufacturers can be of positive assistance in this respect, the customer should be sure to announce promptly to all bidders any changes to the system specifications and to allow them sufficient opportunity to revise their proposals accordingly. The likelihood of

beginning to work almost exclusively with one manufacturer must be avoided at this time for the customer's own sake as well as from the standpoint of being fair to all competitors. As specifications are revised and reestimated one must be on guard against any confusion which is apt to result. Thorough documentation of changes and formal notification to manufacturers will help eliminate possible misunderstandings and future repercussions.

As this process continues, both customer and vendor will wonder sometimes whether the ADP program actually is in the Applications Plan stage or Equipment Selection stage and there will develop an urge to freeze the system. At some point, of course, the system must be frozen in the interest of progressing to the next step. It will be worthwhile, however, to keep the matter open as long as valuable contributions are being made. Avoid the possibility offreezing specifications prematurely to the detriment of management whom the installed system is to serve.

Just as the customer works c losely with manufacturers to validate time and recurring cost estimates for operating the ADP system, he also must thoroughly examine the accuracy of one-time cost estimates such as site preparation, program testing, ancillary equipment needs and parallel operations. These one-time costs can be extremely high and actually can become recurring costs by later having to extend or re-do work improperly planned and estimated.

It is recommended that the controlling factors in selecting equipment be time and cost. Factors which cannot be related to time or cost are difficult to evaluate and their value is easily overestimated, perhaps to the detriment of both the user and the vendors. Unless a given factor can be translated into time or cost, it may later prove to be non-existent or so nebulous as to be unimportant or equal for all qualified vendors.

A major cost consideration, too often overlooked, revolves around the terms of the contract under which the equipment may be procured. Whether the equipment is to be purchased or rented, a complete understanding of the contract provisions offered by each manufacturer being considered is essential in that some of the more important terms affecting costs differ significantly. To assist Government agencies in making an intelligent assessment of rental contracts concluded by the General Services Administration, that Agency has published a Brief of Terms and Conditions which covers the more salient features of each ADP contract and compares them with each other. Being a digest, the Brief should not be used as a sole means for evaluating the contractors' terms but as a device for highlighting differences which then should be studied in detail to determine exactly how they will affect costs of operation in a given installation.

Although GSA contracts do not yet exist for the purchase and maintenance of ADP equipment, prospective purchasers should obtain copies of vendors' proposed purchase and maintenance proposals and subject them similarly to careful scrutiny and comparison. Among the more important considerations in purchase and maintenance contracts will be: warranties; bases for maintenance charges; guarantees for efficient maintenance (preventive and remedial); performance standards; installation guarantees; trade-in values and assurances against inordinate increases in maintenance charges.

These same considerations, somewhat modified, apply to rental contracts, but in addition, rental contracts contain provisions which determine the amount of machine use to which a customer is entitled for the basic monthly rental and the rate charged for usage above that amount. Although all rental provisions are important, the terms of use are the most important and have the greatest effect on continuing costs.

After the final selection of equipment (including approval of the selection), each manufacturer who submitted a proposal should be informed as to whether his equipment was selected. Unsuccessful bidders should not be allowed to place the customer agency in a position of justifying to each manufacturer why his equipment was not selected. On the other hand, the agency should have documented well the reasons for its selection, in the event that information is requested by an authorized Government agency or the Congress.

3. Before Installation

The time between selection and installation of equipment seems to be the period most likely for the honeymoon to end. The ADP marriage of the customer and the manufacturer is in for some rough going unless each party demonstrates his good faith by carrying out his promises. This is the time of detailed development of applications which entails all the efforts required to translate the Applications Plan and equipment selection into a complete and practical operating reality. These efforts include staffing, training, detailed planning, testing and debugging of initial applications, the readying of installation facilities, development of desk procedures, data clean-up, file conversion, etc.

There are responsibilities of the manufacturer such as building the equipment to permit its testing and debugging as components, and again as a system, in time to meet the agreed installation date; providing adequate instruction for programmers and operators; making available sufficient program testing time well before the equipment is installed; providing accurate and firm site preparation requirements insofar as equipment is affected; meeting promises on the availability of canned routines and automatic programming aids; providing technical assistance in program writing, coding, file clean-up and conversion procedures.

The customer also has responsibilities toward the manufacturer such as preparing the site inaccordance with written specifications including an agreed upon date; selecting qualified people to be trained as programmers and operators; completing routines on schedule to permit timely debugging and testing; accepting technical advice from manufacturers' representatives.

In essence, the two parties are now one team with a single objective, which is to implement an effective ADP system on schedule. Representatives of each party must work hand in hand with their opposite numbers. The salesman's work is not finished with receiving the order for equipment. He is now the coordinator of his company's services and works with the customer's ADP manager insuring that specialists are available as needed just as the ADP manager reciprocates the customer's cooperation toward the manufacturer. In order to insure an orderly progression of events, to prevent misunderstanding between the parties and to signal deficiencies before they become deterrents, a timetable should be developed and followed as closely

as possible. This device will serve as a constant reminder of each party's obligation to the other and becomes a check list that each may use to measure his own progress and monitor the other's. Both should bear in mind that failure on the part of one member of the team surely will cause failure on the part of the other. It should be evident that mutual respect and cooperation during this third period is essential.

4. After Installation

Once the equipment is installed and certified ready for use a new relationship develops, that of the ADP operations supervisor and the manufacturer's maintenance engineer. Of course, the older relationships are still present. Training, retraining, technical assistance in optimizing functions and developing new applications are a few of the established relationships that will continue on the same basis as before and which already have been covered. In the continuing operation of the equipment, however, the manufacturer is represented on a "full-time basis" by the engineer whose duty it is to keep the machinery in first class operating condition. His opposite number is the customer's operations supervisor who must have efficient equipment in order to do his job. The necessity for cooperation and mutual confidence is obvious under such circumstances.

The engineer is responsible for performing preventive and remedial maintenance to the equipment. More pointedly, however, his job is to minimize the need for remedial maintenance; an objective in which the customer's employees share responsibility. Downtime due to machine failure can be reduced sharply by adequate checking and testing in accordance with an agreed upon preventive maintenance schedule under which the equipment will be made available to the engineer. Over and above the schedule, the components should be available to the engineer whenever they are not schedule, for use. As a general rule, the machine system should be available for preventive maintenance at least 90 minutes out of each 24-hour period. It is pound foolish to postpone preventive maintenance except in rare emergencies. Therefore, the operations supervisor should make certain that the maintenance schedule is followed carefully.

The maintenance engineer should be notified by the operator as soon as a possible machine malfunction is detected. Together they can more quickly ascertain and eliminate the cause of the trouble. Under no circumstances should the operator attempt to repair an ADP device. Respect for each other's responsibility will help to establish an efficient working relationship. A record of maintenance required should be kept by the operations supervisor in cooperation with the engineer so that if down time should become excessive on a continuing basis, the record will demonstrate the necessity for mechanical replacement of an unreliable machine or device. The same record may be used to note any extra maintenance charges which may accrue, or any periods of downtime which may result in deductions from monthly rental charges. Of course, any necessary action to be taken is a matter for discussion with the manufacturer's sales reporesentative rather than the maintenance engineer.

- -

This session of the Seminar has attempted to outline the importance of the relationships between customer and equipment manufacturer during the several stages of developing and operating an effective ADP system. Navy practices have been cited as a framework for discussing this subject. You may disagree with these practices or you may find the framework inapplicable to your own particular situation. Despite such differences, the session will have served a worthwhile purpose if it has kindled a little better appreciation of the interdependence of customer and equipment vendor in the field of ADP.

GRADUATE SCHOOL
U. S. DEPARTMENT OF AGRICULTURE
SEMINAR ON AUTOMATIC DATA PROCESSING
FOR FEDERAL EXECUTIVES

THE ADP FAMILY AND THEIR GENERAL CHARACTERISTICS

M. H. Hansen and J. L. McPherson

At the 15th International Congress on Hygiene and Demography in a paper relating to tabulation facilities at the Bureau of the Census Mr. H. T. Newcomb said:

> "Too much cannot be said in praise of this machine which has enabled us to compute results with much greater rapidity and accuracy than by the old method of tallying besides giving the opportunity to make a much more thorough analysis of the figures."

Mr. Newcomb did not refer to ADP, probably because he was not familiar with that particular verbal short hand. This unfamiliarity is understandable. The meeting at which Mr. Newcomb made this remark was held in September 1912--over 47 years ago.

Mr. Newcomb was talking about the then relatively new punched card tabulating equipment. He and his coworkers were undoubtedly just as impressed with and optimistic about the potential contribution of punched card tabulating equipment as we, their descendants at the Bureau of the Census, are about electronic data processing equipment. Certainly if it had occurred to them these forebears of ours would have been entitled to call their new devices ADP machines. The equipment invented by Herman Hollerith during the 1880's constituted just as much a break through in his day as the electronic equipment which is the subject of these seminars today.

Many of us now on the staff at the Bureau of the Census can be likened to members of the D. A. R. As we understand them the members of this organization take no small amount of pride in the fact that one or more of their progenitors participated at the birth of our great country. Admittedly, however, none of the present members threw tea in Boston Harbor or pulled an oar at Valley Forge. Similarly none of us at Census punched even a single hole in an 1890 Census punched card or attached a wire to the original Hollerith tabulator. Nevertheless we are proud that Dr. Hollerith was a Census employee during the processing of the 1880 Census where he recognized a pressing need for more automatic data processing equipment than then existed. Like the D. A. R. we enjoy wrapping ourselves in the glory of our predecessors.

Certainly history has justified the enthusiasm typified by Mr. Newton's comment. A giant industry has grown from the invention of punched card equipment. Although we wouldn't know how to go about proving it we believe we can even today safely affirm that throughout the world, more data processing activity is accomplished with punched card equipment than with any other kind of machinery including high speed electronic data processing equipment.

One thesis we want to develop here, for which the foregoing is a brief background, is that ADP is neither new nor has it really been achieved if one interprets the A, for "automatic", at all literally. Data processing means many things to many people. The most nearly completely "automatic" data processors with which we are familiar are devices, which in mid-20th century terminology, are called analog computers. The thermostat on the wall in your living room is a good example of such a device. The data it processes is the temperature. The purpose to be served is the maintenance of a relatively constant temperature. Except for infrequent malfunctions, which can usually be corrected readily, this data processor is truly automatic in the sense that it requires effectively no human attention after the initial parameters of the problem have been stated--i.e. the desired temperature has been communicated to the data processor by setting a dial.

It is, however, the digital or general purpose type of data processing equipment that is the subject of this seminar rather than specialized analog computing equipment. Certainly modern electronic data processing machines make it possible to automize some aspects of data processing to a much greater extent than was possible a decade or so ago. Nevertheless there continues to be a significant amount of work requiring the attention and skills of human beings involved in the successful use of these equipments. From this point of view then there remain many challenging areas in which research and development may bring us still closer and closer to full "automatic" data processing.

We mentioned the invention of the punched card method toward the end of the last century as one bit of evidence that the development of machinery to process data is not exactly new. Most good reference books on modern electronic computers pay homage to Charles Babbage's. As early as 1812 Babbage began work on a device he called a difference engine and from 1833 until his death in 1871 he devoted himself to the design and construction of an "analytical engine." Today competent engineers assure us that Babbage's design was a good one and incorporated much of the logic we find in modern electronic computers. Babbage was unsuccessful because in his day parts could not be machined to the tolerances he specified not because of any defects in his design which incorporated the three main elements which characterize today's computers--namely an arithmetic unit, a control unit and a memory. Incidentally--and we wouldn't be surprised to learn that Hollerith was aware of this--Babbage proposed to use punched cards for input, particularly to the control unit, in his analytical engine.

Thus, we submit, the work of Babbage and of Hollerith as well as the work of others who developed desk adding machines and desk calculators not to mention the oriental who a thousand or more years ago invented the abacus is evidence that the search for devices to "make automatic" in some sense or other the processing of data is not new. What is, in our opinion, new is the machinery that has grown from the marriage of the mathematician, statistician, accountant and others with data processing problems to the electronic engineers who were capable of organizing the products of their industry into the complicated networks of circuits which make the electronic data processor's of today.

Just who deserves credit for first recognizing that the vacuum tubes used in radio transmission and reception had properties which would make them useful devices for computation we are not qualified to say. Dr. Norbert Weiner, of Cybernetics fame pointed out that vacuum tubes could be used to perform binary arithmetic prior to 1940. To the best of our knowledge the first men actually to

assemble equipment electronically to compute were Dr. John Mauchly and J. Presper Eckert at the Moore School of Electrical Engineering at the University of Pennsylvania. During the early 1940s these men developed and built the Eniac (Electronics Numerical Integrator and Calculator). It was built for the Ballistics Research Laboratory at Aberdeen Proving Ground and was completed in 1946. The Eniac although it has general computing ability was a rather special purpose computer. It's main function was to compute firing tables. For this purpose it was eminently successful. It computed a trajectory in about half the time it takes a shell to travel from the gun to the target.

While they were still building the Eniac which was during World War II Eckert and Mauchly recognized the potential value of electronic computers for general data processing purposes. They began to formulate, at least in their minds, the design of a general purpose electronic computer. Dr. Mauchly's main interests were in astronomy and physics. Mr. Eckert was an electronic engineer. Neither of them was particularly familiar with the problem of accountants and economists and social statisticians and others with large data handling responsibilities not in the domains of physics and engineering. They sought the advice and consultation of persons with experience in mass data handling and among other organizations they visited during the early 1940's was our Bureau of the Census.

Their early visits were frankly for the purpose of finding out the kinds of problems we faced and understanding what machinery we were using to solve them. It was generally understood that their endeavors in the area of electronic computing machinery would have to be devoted to equipment to facilitate our war effort and that not until hostilities ended would they be able to devote any important share of their time to the development of equipment for civilian use. During this period our attitude at the Census was a peculiar combination of awe and doubt. On the one hand it was obvious that these gentlemen were respectable citizens, presumably responsible, since they had a contract through a major university to build equipment for the United States Army. On the other hand they seemed to have the crazy notion that somehow we could use our nine tube superhetrodyne radio receiver to tabulate census results. We willingly answered their questions about our work and showed them our punched card processing equipment. We now fear, however, that in those days we classified them as just a couple of sightseers.

When the war ended, however, Eckert and Mauchly made serious representations to us about building electronic data processing machinery for our use. It then became important for us to decide just how seriously their proposals should be taken. There existed in the mid 1940's a "Science Committee" composed of representatives of the various bureaus which make up our Department of Commerce. At our request this committee met to hear us describe the proposals we had received from Eckert and Mauchly and to advise us on a way to proceed to evaluate the validity of their ideas and their competence to implement them. It was the recommendation of the committee that scientists at our sister agency in the department, the National Bureau of Standards, could be of real help in understanding and evaluating the technical aspects of the proposed electronic computing equipment. Dr. Archibald McPherson, Deputy Director of NBS was a member of the Science Committee and played a major role in initiating a cooperative arrangement between our Bureau of the Census and the NBS which continues to this day.

After preliminary meetings at which we familiarized, as best we could, the mathematicians and engineers at NBS with the general character of our work it was arranged for Eckert and Mauchly to describe their proposed equipment to a group of NBS scientists. We were pleased and relieved when our colleagues at NBS told us that we were not being duped by a couple of charlatans who wanted to sell us the Emperor's new clothes but that electronic computing in the Eckert and Mauchly way was indeed a real possibility and that these two inventors had made a very good impression at NBS.

In 1946 we transferred $300,000 of our funds at Census to NBS to defray the cost of work they were doing in our behalf. Here as an aside we might report how the fates and the Congress of the United States were good to us. As you know our war effort had moved large numbers of people in our country; many people had migrated from one area to another to make their contribution at defense plants. It was decided that a population sample census to be enumerated in 1946 should be taken and Congress had earmarked in our fiscal 1946 appropriation funds to defray the costs of getting ready for this Census. When it came time to pass our fiscal 1947 budget, however, Congress changed its mind about the desirability of a population sample census and did not appropriate funds to conduct it. Fortunately, however, they did not impound the get-ready money they had supplied the previous year and we had no difficulty in getting their permission to transfer $300,000 to NBS.

Some of this money was used by NBS to pay for a study contract they awarded to Eckert and Mauchly in 1946. By this time these men had left the University of Pennsylvania and established their own organization. One result of this study contract was a general design of the computer now known as Univac I.

While the study contract was in being other organizations particularly the Air Forces, the Army Map Service and the Office of Naval Research developed interest in the possibility of electronic computation and sought the advice and assistance of the NBS personnel with whom we were working. These defense groups were able to finance work on a much larger scale at NBS than we were able to underwrite. A large group of competent scientists were assigned responsibility at NBS in this new and interesting field.

In June of 1948 the National Bureau of Standards entered a contract with the Eckert-Mauchly Co. for the construction of a Univac for Census. Shortly thereafter the contract was supplemented with an order for two additional Univac systems one for the Air Comptroller and one for the Army Map Service.

The original delivery date specified was February 1, 1950. The first Univac ever built was actually accepted by the government on March 31, 1951. Incidentally it is still in use today at our Bureau where it is performing quite satisfactorily. In 1954 we acquired a second UNivac I and last year we added two Univac 1105 systems at our Washington offices and made cooperative arrangements for the installation of compatible 1105 systems at the University of North Carolina and at the Armour Research Foundation of the Illinois Institute of Technology. At each of these institutions we will have first priority on from one-half to three-fourths of the computer time during the next two years which spaces the period of our peak requirements for processing the 1960 Decennial Censuses of Population and Housing.

The period from 1948 to 1951 during which the government had contracts for general purpose digital computers but no hardware supplied by private enterprise was certainly stimulating and some-times trying. The Congress maintained a lively interest and followed the development closely. Great pressure was exerted on NBS personnel by our Bureau of the Census, by the Air Comptroller, by the Army Map Service and by the Congress to "speed up" the production of the equipment although in retrospect we all now recognize that just how they should go about effecting a speed up was an administrative detail we were quite willing to leave up to them. They on their part worked with a will and although they didn't find any magic formula to make contractors deliver on schedule they accomplished a near miracle by designing and building SEAC. This computer actually began producing useful computations on May 9, 1950 almost a year before the first Univac was accepted. This was a real accomplishment particularly in view of the fact that this endeavor began at NBS after the Univac contracts had been awarded.

During the early 1950's many organizations entered the electronic computer field. Most notable was IBM. The energy and enthusiasm with which they worked in this area once they had decided to enter it was truly admirable.

The first machine in the 700 series, the 701 was delivered in 1953. Thereafter its successors came along in rapid succession: the 702 and 704 in 1954, the 705 in 1955, the 709 in 1958 and just in the past year they are offering the solid state 7070 and 7090.

GRADUATE SCHOOL
U. S. DEPARTMENT OF AGRICULTURE
SEMINAR ON AUTOMATIC DATA PROCESSING
FOR FEDERAL EXECUTIVES

THE PRINCIPLES OF PROGRAMMING

By Ethel Marden

Introduction

It is not possible to cover the subject of programming within the space limits allowed here. However, there are available several recent books which give more detailed treatment to this topic. 1/ In addition, there are several periodicals which report current activities in the computer area, including programming. Perhaps the best such publication for keeping abreast of the new developments (which are occurring at an extremely rapid rate) is the monthly magazine published by the Association for Computing Machinery (ACM). 2/

Programming, in the sense intended here, is the art of communication with a computer. Since at this time the language of humans differs from that of computers, programming implies communication with the computer, either in the computer's own language or in an intermediate language which can be understood by computers. It may be that in the not distant future, we will have better educated computers which can speak English, or other languages employed by humans, but that situation does not prevail now. Therefore, learning "programming" means learning to communicate with the computer in some language which will be understood by automata, or by an intermediate "interpreter." The Association for Computing Machinery published a glossary of programming terminology in 1954, 3/ and in it they define "program" as follows:

> program (noun) - a plan for the solution of a problem. A complete program includes plans for the transcription of data, coding for the computer and plans for the absorption of the results into the system. The list of coded instructions is called a routine.....
>
> program (verb) - to plan a computation or process from the asking of a question to the delivery of the results, including the integration of the operation into an existing system. Thus programming consists of planning and coding, including numerical analysis, systems analysis, specification of printing formats, and any other functions necessary to the integration of a computer in a system.

Development of Computer Languages

Communication implies the use of a language, and a language is usually made up of discrete words; computer languages are no exception. The unit of a computer language is usually a "word," which is fixed in length for any one computer but which varies in length among different computers. Some may be as short as 16 binary bits*; most modern ones contain from 60 to 70 bits. Some computers are designed for variable length words, where the programmer chooses for each word as few or as many bits as he needs, and then inserts a marker, which may be thought of as a space, to indicate the termination of the word. Such words written for a computer may consist of instructions, or commands, given to the computer to execute; on the other hand, the words may represent numbers or data in other forms, which the computer is required to manipulate in carrying out the series of commands issued to it by means of the routine (or program).

The marvelous mechanism which is a modern electronic computer is based on the simple notion of there being two electronic states, one represented by a "0" and the other represented by a "1" in the binary numbering system. From that beginning, within a short period of time, complex communication systems made up of many languages have evolved. It is necessary to understand something about the binary numbering system in order truly to understand what programming is all about.

Binary Numbers

Binary	Decimal
0	0
1	1
10	2
11	3
100	4
101	5
110	6
111	7
1000	8
1001	9
1010	10
1011	11
1100	12
1101	13
1110	14
1111	15
10000	16
- -	-
- -	-
100000	32

*A "bit" refers to a "1" or a "0" in the binary system of notation.

It is easy to see that addition of successive powers of 2 results in the accumulation of additional digits to the left (in the case of integers), as in the decimal system, where the decimal number:

12358	equals		8 x 1	8
		plus	5 x 10	50
		"	3 x 100 (or 10^2)	300
		"	2 x 1000 (or 10^3)	2000
		"	1 x 10,000 (or 10^4)	10000
				12358

just as, in the binary number system, the binary number:

11010111 is equal to:		binary	decimal
	1 x 1	1	1
plus	1 x 2	10	2
"	1 x 4 (or 2^2)	100	4
"	0 x 8 (or 2^3)	0000	0
"	1 x 16 (or 2^4)	10000	16
"	0 x 32 (or 2^5)	000000	0
"	1 x 64 (or 2^6)	1000000	64
"	1 x 128 (or 2^7)	10000000	128
		11010111	215

The same relationship holds true for fractional numbers, except that successively increasing digits past the decimal point move to the right, (with successively smaller powers of 10). For example, the decimal number:

.98703 is equal to:		9 x .1 (or 1/10)	.9
	plus	8 x .01 (or $1/10^2$)	.08
	"	7 x .001 (or $1/10^3$)	.007
	"	0 x .0001 (or $1/10^4$)	.0000
	"	3 x .00001 (or $1/10^5$)	.00003
			.98703

and the binary number:

.1101 is equal to:		binary	decimal
	1 x 1/2	.1	.5
	1 x 1/4 (or $1/2^2$)	.01	.25
	0 x 1/8 (or $1/2^3$)	.000	.000
	1 x 1/16 (or $1/2^4$)	.0001	.0625
		.1101	.8125

All numbers given to the computer to work with, all data, whether natural language text, symbols, combinations of numbers, language and symbols (known as "alphanumeric" information), and all instructions or commands, must be entered into the computer in the binary numbering system. This does not mean that the programmer must write all of his instructions by means of only two digits; the peripheral or auxiliary equipment connected with the computer takes care of breaking down more complex characters into their basic components, so that programs can be written by means of numbers alone, written in natural language, in algebraic symbols, or in combinations of all of the above. Provision is either made for their reduction to binary form by means of special computer components, or, more usually, by means of special programs written for the purpose of translating simply written programming steps into the more complex (for humans) language of the machines. There is an entire category of special programs for the compiling, assembling, and translation of these simpler programs into machine language. Preparing the class of manipulative programs partially defined above is known as "automatic programming."

Just a decade ago, the SEAC, the forerunner of modern computers, was dedicated at the National Bureau of Standards. As we look backward, it truly seems fantastic that there has been the extraordinary advance in the design of computers that has occurred in this ten-year period. Yet as great as that advance has been, there has been a corresponding advance in thinking with regard to the means of communication with computers: programming. Early programs were written in the machine language which the computer could understand. It was laborious to write and debug * such programs, even though the binary expressions were written in number bases other than binary: octal or hexadecimal or, in some instances, in decimal. Programming then meant communicating with a computer by writing in an arbitrary machine language for that particular computer a series of detailed instructions. The instructions called for the execution of a plan; usually the plan represented the steps necessary for the solution of a scientific or mathematical problem. That type of problem made up the major part of a computer's workload in the dawn of computer history. In an effort to shorten programming time and make it easier to add modifications, early programmers resorted to "relative programming" or "relative coding," where steps were programmed and placed in relative position to other steps with respect to their locations in the computer's memory, or storage space. The fixed sequence of steps was then determined when the program was complete, usually by the assistance of another program written especially for the purpose, and permanent locations were then assigned to each instruction and each data word in the program. This type of procedure resulted in the birth of one type of automatic program.

---

*"debug" is an art term which means to detect and remove the mistakes and errors which are to be found initially in virtually all computer programs.

Nowadays types of problems encountered in an electronic computer installation fall into many different categories; there is a consequent effect on the determination of programming procedures for the varying types. In addition to the scientific problems, there is the large class of so-called "data processing" problems, into which fall all of the business and accounting types of problems, the large inventory and stock control problems, management decision making problems, information retrieval problems, which again include several sub-classes, such as literature searching and machine translation problems, and many others.

The category of programming which is generally considered to be the most difficult is that of automatic programming, which includes that entire class of programs which operate on other programs as data. They rearrange, compile, assemble, translate, or otherwise process hand-written programs according to a preconceived plan, in order to assist the programmer and make his task of communication with the computer a simpler and easier one, as well as to shorten the preparation time of programs. One member of this class of programs is called an "interpretive" program. An interpretive program examines instructions which have been written for the computer, and executes the instruction in each case in a particular manner, which may vary from instruction to instruction, in accordance with stored symbols which make up a part of the word, or instruction, or pseudo-instruction being examined. The interpretive routine may carry out yet another kind of operation, or operate on the instruction itself, depending upon the pre-stored directions which make up the routine.

From the early beginnings of the development of automatic programs, there have been many advancements, and each attempts either to shorten the time for writing a program and/or to ease the burdens of the programmer. For the average program, as much money goes into the writing and creation of the program as into the rental for the hardware which is required to execute it. Programming is expensive at best; if there can be ways of shortening the time for writing programs, or of making it possible to write programs which can be debugged more easily and quickly, then automatic programming is economically very worthwhile. The high cost of programming has been a headache to almost every computer installation in the country--and to those who can say, "The money be damned," the consumption of time required to write programs is of grave importance.

The first sizeable advance in the evolution of the language of programming was the creation of a symbolic language, with compilers to interpret the pseudo-language and translate and assemble it into the language which the computer understands. People do not think in the patterns according to which computers operate, nor do computer operations follow our thought processes--at least not at this stage of their development. It is too rigorous for the human to confine thought processes to expression in the rigid formats of computer language. In addition, it is a tedious and boring and lengthy procedure--hence more errors are likely to be introduced into programs. When it became possible to write orders to the computer in a symbolic language, some fetters were removed from thinking, and there was

greater freedom to consider the steps to be taken in obtaining the correct, or a satisfactory, solution; this procedure is preferable to devoting all of a human's faculties to the expression of each precise instruction in the rigid computer code.

Just as there has been a steady evolution of written and spoken language since the early pictographs in caves, and the later hieroglyphs, so has there been an evolution in the language of programming, but it has been encapsulated into a much smaller time span. And with the changes which have occurred in the last ten years before our eyes, we can expect even greater changes in the next ten. Because programming any sizeable problem is an arduous and tedious task at best, and tends tocheck the free range of imagination because of the mechanics of accomplishing the task, there has been thought given not only to devising simpler languages, but on the other hand to creating more powerful and more sophisticated languages at the same time. A group of people, primarily at the RAND Corporation, have felt that more complex problems can be handled, particularly of the heuristic type, when more powerful tools in the way of programming languages have been developed. 4/ Their thought trends lie in the direction of creating both more complex languages and more complex machines with which to communicate in those languages. They have called them "information processing languages" (IPL), and they are now in the process of developing IPL VII.

From the first direct machine languages, there has been a progressive development of increasingly complex languages in order to make man-machine communication easier. Development of symbolic languages led to the creation of some of the new special languages, some of which are oriented to the statements of algebra and other mathematical languages for dealing with scientific problems. One of the newest of these is the international language known as "ALGOL" or Algorithm Language; another new special language is called "COBOL", for Common Business Oriented Language; the first short-range COBOL has just been completed. Both intermediate and long-range projects for business oriented languages are under way. Six manufacturers are cooperating with government agencies in their development. A problem written in COBOL can be run on any of the computers cooperating in its development, by means of a specially written translation program from COBOL to the computer of interest, and each of the six manufacturers has completed, or is in process of completing, such languages.

In spite of all of the new developments which have taken place in automatic programming and in the creation of new computer languages, there is still need for others in some of the areas where there has been more recent application to computers of the subject areas. This is particularly true in the information retrieval field, which embraces machine translation, document retrieval, information retrieval for content or informative content of a document, and character recognition researches.

Effect of Computer Design on Programming

Relative difficulty of programming among computers is affected strongly by the characteristics of computer design, which differs radically in some cases from

one computer to another. Some computers have single address codes, and these may be likened to the design of a desk calculator which has a register visible to the operator, where results are accumulated. A single-address computer, which can refer only to a single memory address in any one instruction, has such a register, which is normally called an "accumulator." Here, if two numbers are to be added, one instruction is necessary to place one of the numbers in the accumulator; another instruction is required to add the second number to the first, and yet another instruction is given to store the result in a designated memory location. For example, in the IBM 704, this sequence of instructions might read as follows:

> Let address 4000 contain the first number,
> " " 4001 " " second number,
> and let the answer be stored in
> address 6000
> (This naturally presumes a memory capacity
> of more than 6,000 locations)

Instruction	Address	Explanation
CLA	4000	Clear accumulator; put 1st number in accumulator
ADD	4001	Add 2nd number to 1st; accumulator contains the sum
STO	6000	Store the sum of the two numbers in address 6000 of the 704's memory

It will be noted that only one memory location is referred to in any one instruction.

A three-address computer, on the other hand, would be able to accomplish all of the above with a single instruction. For example, the above series of instructions in the NBS SEAC would read as follows:

Instruction	1st Add.	2nd Add.	3rd Add.	Explanation
ADD	4000	4001	6000	Add contents of address 4000 to 4001 and store in 6000

There are advantages and disadvantages to each of the two kinds of computers noted. There are also two-address computers and four-address computers, although

there seems to be a tendency toward the design of more one-address computers than any other kind at the present time.

From the above examples it might seem that programming is a very simple procedure; that is not the case. The simplicity arises in the nature of the example chosen, since it is very easy for either humans or computers to add together only two numbers.

<u>Illustrative Sequence of Programming Steps</u>

The steps necessary to complete a program usually fall into the following categories:

1. Problem definition, including
   a. Numerical analysis, if required, in scientific problems, and determination of most desirable procedures
   b. Organization of project into general logical sections, if possible, in the case of data processing problems; systems analysis

2. Preparation of flow chart

   a. General flow chart, easy to read by systems analysis, management personnel, and others not necessarily familiar with computer-oriented techniques.
   b. Detailed flow charts, broken into sections, from which programmers will work to write code

3. Writing of code (usually in symbolic form, or in FORTRAN, ALGOL, COBOL, or others of the special languages now employed)

4. Preliminary general manual checking of code for organization, completeness, etc.; punching of cards or tapes for program and data; and compiling or assembling of code for its debugging runs

5. Debugging of code

6. Planning for production runs
   a. Further organization of data handling, as required
   b. Changing or streamlining code for greater efficiency in production runs, if applicable
   c. Orderly dispersal of results of computer runs

7. Preparation of necessary reports
   a. Operating instructions
   b. Description of procedures employed in program
   c. Evaluation of results being obtained from computer

d. Recommendations in the light of operating experience
   (1) Concerning modifications to present procedures
   (2) Further mechanization of related or dissimilar procedures
e. Other desirable documentation

After programs are completed so far as the preliminary code is concerned, the building in of any error checks which are possible is almost a necessity. There are certain natural characteristics inherent in data which lend themselves readily to such checks, but such checks must be introduced by the programmers, frequently with assistance on the part of subject matter specialists who may have greater insight into the nature of the information being handled. For the price of a few additional operations, many mathematical checks are available and are not particularly time-consuming. Such checks are well worth the small amount of time, space and effort expended on them. In business or information retrieval types of problems there are usually certain redundancies which lend themselves to checks. In other cases, repetitions may be of value, such as the use of a different route or different procedure to obtain the same results, where "All roads lead to Rome."

Difficulties in Programming

One of the principal difficulties encountered in programming is in obtaining a correct and complete definition of the problem to be programmed. Most scientific problems which can be expressed in terms of mathematical equations lend themselves easily to definition. And a problem which can be defined can usually be programmed easily and in a straightforward manner. Mathematical notation has been developing for several hundred years, and such problems employing it can be expressed succinctly and precisely. That is not true with most data processing problems. Here the descriptive language has not crystallized; there are discrete and singular situations to take account of, as well as many-branched trees and hydra-headed monsters. There are diffuse situations and anomalies which must be clarified and dealt with satisfactorily. There is a mass of detail: each item in the detail must be recognized and its place in the structure preserved, under varying conditions, as long as required. The problem of organization for most data processing problems is a formidable one, including as it does the diverse input data, sometimes from many sources, the intermediate results, usually with some cross-correlation involved in the processing, and the final results and end products.

Although it may be considered difficult to formulate a scientific problem so that there is nothing whatever left out, it is next to impossible to define any one of the large data processing problems with which we are continually faced today. They usually grow, like Topsy, as they are put into production by teams, and omissions and additions are tacked on wherever possible until the point of diminishing returns is reached and the entire structure must be replanned, reprogrammed and recoded to include all of the improvements and modifications. The problem definition, when a

plan is all-encompassing, is so difficult (even for subject matter specialists who have lived with the material making them up for years) because the procedures of the problem have usually never been completely documented. It is difficult for people in such areas to call to mind all of the situations which are usual and which they deal with daily; it is even more difficult to call to mind situations which are not usual, but which they might be able to cope with when or if they arose. The computer, on the other hand, cannot do on-the-spot improvisation, unless it has been previously directed in detail as to what steps to take in each instance, and of course many could never be foreseen. Again, most subject matter specialists know things which they do not know that they know, and it requires experience and prodding to elicit all of the information concerning a problem from them; they are like sponges, which give up information (as sponges do water) over along period of time, when pressed. Yet complete information is essential, because every detail must be given the computer in advance of operation, and those details must be written into the computer program. Every decision must be anticipated; the one thing which computers cannot do yet is to think--I say "yet", because there is evidence that they will be simulating thought processes within the next few years, if promising advances in research indeed do develop.

Training and Aptitude

The selection and training of programming personnel is an extremely important function of administrators. What qualifications are required of potential programmers? For scientific and mathematical programming, a knowledge of mathematics is of course an essential. For certain categories of problems, physicists and statisticians possess an adequate background so far as theoretical training is concerned. From time to time there are trends of thought which say, "Take any clerk and you can make him into a good programmer." That is simply not true, although there have certainly been some exceptions. Some knowledge of mathematics is usually helpful, even when mathematical equations are not being handled, because of the training in abstract thought which mathematicians undergo.

There are in general two schools of thought with regard to selection of programming personnel.

1. One school says, "Employ trained programmers from the beginning; have them work with subject matter specialists and let them learn the subject matter as they go along."

2. The other school says, "Use subject matter specialists and teach them programming."

There is something to be said for each school, of course, and perhaps there is no pat answer to fit every situation. The answer certainly varies with the nature of the subject matter. Certainly not all subject matter specialists can become good

programmers, despite either their intelligence or how well-grounded they are in their profession. In certain areas, because of the length of time required to gain knowledge of the subject matter, it is highly desirable to have subject matter specialists learn programming. In certain specialized areas, it may require years to give a trained programmer a thorough background in the particular subject area in order to make him competent in both skills. Frequently, the solution is to bring together teams in which subject matter specialists work alongside programmers for a highly efficient operation. In information retrieval problems, linguists, librarians, researchers and those possessing similar skills are certainly better versed in techniques of literature searching by manual methods than those with no similar experience. Qualifications for programmers in these areas are somewhat different from those heretofore required for scientific and business data processing problems, from the subject matter point of view.

As in most professions, those with an aptitude for programming learn more quickly and are more ingenious and efficient programmers after the learning period is over. Certainly almost anyone can learn programming, but it is not economically feasible to have inefficient programmers in an installation. Costs of programming can easily be doubled by unwise selections of programmers; they are certainly greater when inefficient ones are maintained in an installation.

Certainly the first criterion I would list for selection of programmers is that of above-average intelligence. An ability for deep concentration is required in programming, as are imagination, a tendency to unconventional or non-stereotyped thinking, and flexibility and adaptability, particularly in the light of changing conditions and new developments in this rapidly changing field of technology. One good single criterion for a programmer is the liking or the ability for solving puzzles.

The importance of well-organized training courses cannot be overemphasized. Manufacturers' training courses are not sufficient; they need to be supplemented by on-the-job training courses, with exercises slanted toward the application in that particular laboratory. After formal training courses are completed, it is usually preferable to require neophytes to work in an apprentice relationship with more experienced programmers for the first two or three trials, perhaps switching programmers at the end of each such trial. This gives the amateur an opportunity to observe techniques employed in working situations as opposed to illustrative ones usually studied in class. If he works with two or three different senior programmers, he is also able to observe how the approach of one differs from that of another, so that he does not get "set in a mold."

Programming is fascinating: that statement is partially true because programming is a demanding and an absorbing profession. And when there is a reasonable assurance that the program is "debugged" and can now be used for production, the programmer feels somewhat akin to Christian in Pilgrim's Progress, when he has finally surmounted all of the difficulties with which his journey is fraught. He has arrived at last in Heaven.

Bibliography for

The Principles of Programming

by Ethel Marden

1. Digital Computer Programming, by D. D. McCracken, John Wiley & Sons, Inc. New York; Chapman & Hall, Ltd., London. 1957.

2. Communications of the ACM, Association for Computing Machinery, 2 East 63rd Street, New York 21, New York.

3. First Glossary of Programming Terminology (published by ACM as a Report to the Association for Computing Machinery from the Committee on Nomenclature, June 1954), Association for Computing Machinery, 2 East 63rd Street, New York 21, New York.

4. "A Command Structure for Complex Information Processing" by J. C. Shaw, A. Newell, H. A. Simon, and T. O. Ellis, Proceedings of the Western Joint Computer Conference, Los Angeles, May 6-8, 1958. Also: New York, American Institute of Electrical Engineers, March 1959. p. 119-128. Also pub. by RAND Corp., Santa Monica, Calif. (RAND P-1277).

GRADUATE SCHOOL
U. S. DEPARTMENT OF AGRICULTURE
SEMINAR ON AUTOMATIC DATA PROCESSING
FOR FEDERAL EXECUTIVES

INFORMATION RETRIEVAL

By Ethel Marden

Information retrieval is generally grouped into three principal categories: (1) literature searching, which is also known as "information retrieval", which may be divided into (a) retrieval of documents, or (b) retrieval of informative content of documents; (2) machine translation (MT), and (3) character recognition. Also under the general heading of information retrieval there are loosely included such topics as self-organizing systems and artificial intelligence; the latter topic again splits to include such work as the simulation of neuron networks and heuristic experiments such as theorem proving operations, chess playing programs, etc.

Automation of such procedures as those listed above must perforce be accomplished by the aid of computers, in the absence of more suitable equipment for those applications. Until appropriate "symbol manipulators" are designed for the several diverse categories making up the continually broadening class of operations now designated as "information retrieval", they must be simulated by the one adequate class of automata readily available to most research laboratories: computers.

The major part of this discussion will be confined to the first of the principal categories listed above.

Character Recognition

Research on character recognition procedures is proceeding simultaneously in several organizations, but the final objectives to be attained are diverse. Banks have a strong motivation for development of character recognition, and have contributed heavily to research which will benefit their specialized needs. Research on the use of magnetic inks for individual check recognition has made great progress under such sponsorship. The Bank of America underwrote the development of ERMA, a specialized computer designed by Stanford Research Institute for automatic processing of check handling procedures, which was put into operation in the San Jose area in early 1956.[1/] Coincidental with that development, the Research Institute of the National Association of Bank Auditors and Comptrollers (NARAC) and the American Bankers Association were sponsoring research along similar lines in the use of magnetic inks in character recognition for banking purposes.[2/], [3/]

For several years, the U. S. Post Office Department has been conducting research programs leading to a greater mechanization of postal procedures,

particularly in the areas of letter sorting and direction.[4]

More advanced evolution of character recognition procedures are awaited for applications in both the areas of machine translation and in literature searching; such assistance in solution of their complex problems is badly needed. Some valuable ground-laying work was financed by military contracts, and later work has benefitted by the base so developed by the pioneers. Study of the problems includes the capacity for recognition of alphabetic characters, whether script or printed, and, if the latter, no matter what the type font. It has been demonstrated that even poorly made characters, or characters imperfectly reproduced because of irregular transfer of ink or other causes, can be recognized in some cases after analysis of the outlines of the figure itself. This analytical operation also has the capability of reducing the amount of "noise" which may be present in or around the character being analyzed.[5] A certain amount of development in this field has made available some commercial machines for this kind of activity.[6]

For certain information retrieval problems, the ability to distinguish between almost identical drawings or diagrams, or to recognize portions of diagrams contained in larger configurations is of importance. Recognition of chemical structures is an important concomitant of the latter case. These areas of research are of great interest to those concerned with searching engineering drawings, or with making patent searches. The exercise of a small amount of imagination enables one to visualize the startling results which a major break-through in this region will bring to bear on all clerical tasks.

Machine Translation (MT)

The primary interest in automatic translation in this country at this time is in the translation from Russian to English, and that is where the greatest effort is concentrated. It is estimated that 8 per cent of the world's scientific journals are published in Russian, and that percentage is expected to rise. It is generally felt that it is more important that this country be in a position to learn of important scientific research being accomplished in Russia (and therefore reported in results published in Russian) than in any other country whose language differs from our own. Until the last year or so, current results of Russian research were largely unknown in this country. In some instances, valuable time was thus lost in applying such results; again, there was duplication in many instances of research conducted elsewhere because of the inability to obtain translation and dissemination of results of foreign work. There are relatively few people in this country who read Russian, and consequently not enough to translate quickly the scientific documents which are in the open literature. In addition, human translation is both costly and slow. It is estimated that a human translator's output is about 2,600 words a day; in apposition, modern computers have an output in MT of about 20,000 words per hour, and it is expected that computers of newer design and now coming into operation will translate at a much faster rate.

A recent Congressional committee instituted investigations concerning the prodigious amounts of money which have been spent in the last few years on MT, in an effort to determine whether the results justify the enormous expenditure of funds (mostly government) which has been made. Money has been poured into this endeavor on many fronts, in a crash program. Most people are of the opinion that the research, however costly, is worthwhile, and recent indications are that it is paying off. There are several installations which are able to produce rough translations at the present time, which require a small amount of editing by human translators. There is great promise that in the relatively near future the major part of this editing may also be accomplished by the machines, and that they may improve upon the performance of humans, except perhaps in specialized areas where the terminology has not yet crystallized.

It may be of interest to note the effort which Russia is expending in MT. In 1955, there were two institutes active in research on MT: The Institute of Precision Mechanics and Computer Engineering headed by Dr. Panov and The Steklov Institute of Mathematics, headed by A. A. Lyapunov. Both groups employed about 12 scientists. There are now 79 different institutions engaged in research on MT, consisting of approximately 500 mathematicians, linguists, and clerical personnel.[7]

Along with the concentrated interest in MT, there is an attendant acceleration of interest in linguistics. Indeed, study of linguistics forms a necessary part of any MT operation. One prominent linguist remarked the other day that for the first time, a linguist is able to earn a living in his chosen profession. As in other areas, the approach to the study of linguistics varies with the individual. Noam Chomsky, of MIT, holds the belief that any phrase structure grammar has to be put through a transformational grammar in order to adapt it to MT. Victor Yngve, a former pupil of Chomsky's, says that English is "left to right" and that a transformational grammar is not required for MT. Zelig Harris, at the University of Pennsylvania, is using an algorithmic system which is quasi-mathematical in nature. Ida Rhodes, of the National Bureau of Standards, is an advocate of the so-called "push down" system. Other groups who are engaged in a major effort in MT are concentrated at The RAND Corporation and at Ramo-Wooldridge Corporation, both in the Los Angeles area, and at Georgetown University. It is believed that there will be widespread applications in other areas of information retrieval of the effort expended in linguistics for purposes of MT.

<u>Information Retrieval and Literature Searching</u>

<u>American Documentation</u> makes a good attempt at covering the most important work going on in the area of mechanical literature searching. These projects range all the way from small punched-card operations to very large-scale computer projects. Recognition has been accorded this kind of information retrieval work by the major computer conferences each year. The annual ACM conference, as well as the EJCC and WJCC conferences devote special sessions to papers dealing with information retrieval subjects. In addition, two international conferences have recently been held exclusively on the problems of

information retrieval; the International Conference on Scientific Information (ICSI), held in Washington, D. C., November 16-21, 1958, and the International Conference on Information Processing (ICIP), Paris (FRANCE), June, 1959. In addition, each year there are several ad hoc conferences, or symposia, devoted to the same subject matter.

The terminology among information retrieval workers has not yet been standardized; this leads to some difficulty in appraising the content of the literature which has been published on the subject. Several projects have attempted to do something about this lack, among them the American Standards Association, the Western Reserve University Conference Committee, and the UNESCO and IFIPS groups interested in information retrieval.

Recognition of the need for improved utilization of scientific information prompted the establishment, on November 17, 1958, of a Research Information Center and Advisory Service on Information Processing, to be operated jointly by the National Science Foundation and the National Bureau of Standards, and which has received financial assistance from those two organizations and in addition from the Council on Library Resources. Recent hearings before the Committee on Government Operations of the United States Senate, under the chairmanship of Senator McClellan, pinpoint the problems and needs in this area very forcibly.8/

Information retrieval lends itself to many different kinds of divisions, one of which is the clear break between library-oriented and non-library-oriented applications. Librarians have recognized a need for modernization of library procedures, but have in some instances been more conservative than others in their approach to the use of computers for improving techniques. Chemists, perhaps more than any other class, have been concerned with the need for mechanical aids to literature searching. They have evidenced their interest by the publicity given over the years in their literature to the efforts made in the notation, indexing, and abstracting areas and, in later years, to the various projects for literature searching by means of punched card machines, special purpose machines of several categories, and computers. Perhaps workers in the chemical area have been made aware of the need for mechanical aids because the chemical literature has been increasing at a more rapid pace than most of the other technological literature, in an era which has seen an explosion of technical publications in all fields.

Again, it may be of interest to note the Russian activity in this area. The All-Union Institute of Scientific and Technical Information is reported to have some 2300 full-time and 20,000 part-time workers engaged in abstracting technical and scientific literature, both Russian and the literature from other countries.

Problems encountered in setting up systems for mechanical literature searching range from those of a linguistic nature to considerations concerning requisite equipment, with many intermediate stages such as the indexing and abstracting questions, and the design of systems to accomplish the desired objectives. The problems are indeed formidable, and almost every literature searching project now in development is encumbered with some which appear to be unresolvable at the present time. However, not everyone takes so pessimistic a view as that expressed by Yehoshua Bar-Hillel: 9/

> "All attempts made so far at establishing a general, mathematical theory of literature searching must be regarded as failures........The fashionable appeal to lattice theory, symbolic logic, semantics, and thesauri is shown to be unsubstantiated. It is proposed to investigate literature search systems in which the role of the computer would be merely that of performing utterly routine and (for humans) time-consuming operations, under constant control of the user."

Studies are under way by groups engaged in special efforts on resolution of difficulties which represent in many cases segments of the main problem. One primary difficulty embraces the entire area of linguistic considerations. Language in the form of discrete words is the usual means of communication in printed or written material. Language in this form is not understood yet by computers without an intermediary of some kind. A synthetic solution of the difficulty of communication between humans and computers is that of creating an intermediate language of symbols, where a unique symbol expresses a word or an idea, and the computer is programmed to substitute the symbol for the idea, or "thing", so expressed, and to recognize the concept or the object by means of the symbol expressing it. This sort of procedure, however, is practicable only in files of limited subject matter and size. Further complications are introduced with the passage of time, such as semantic changes, as well as changes in the accepted values of such things as physical constants, accepted values of precision in measurement, or even in the standards of measurement employed.

E. de Grolier has proposed the creation of a general machine language, which would also be a link to join different translating machines working with natural languages.10/ He points out, however, that many scientific disciplines must be brought together for this work, including classification, linguistics, and logic. V. Yngve, at MIT, advises the use of an automatic programming system as an intermediary between the linguist and the computer, where the linguist employs in his research a notation or language called "COMIT"; that language in turn is converted by a conversion routine or compiler. Through such a language, Yngve states that the linguist can direct the computer to analyze, synthesize, or translate sentences. He advocates the use of English text in literature searching because of the possibility of missing a reference when indexes or abstracts are used, in spite of the complexity which this procedure introduces into the

searching.[11] Thyllis Williams, at Itek Corporation, maintains that the informative content of a document is usually represented incompletely in indexing it, and that under her system of "L-indexing", it can be developed to a greater depth and specificity for a searching system employing selective indexing.[12] "L-representation" applies to natural language a "normalized representation" for the storage of index data, and Miss Williams employs an "L-dictionary" of special rules for correlating elements of non-normalized expressions.

Difficulty of developing a system for literature searching varies, of course, from project to project, and depends partially upon the objective of the search, the magnitude of the file or library which must be searched, and the fineness of discrimination which must be built into the search program. The scale of the search might run from a simple hunt for a matching name or title to an all-encompassing search similar to those required by the U. S. Patent Office.

While it is likely that one could develop a matematical model of any existing retrieval system, no one has yet set forth a comprehensive model which might be generally applicable. A number of research workers have concerned themselves with developing general theories in this field, and the recent International Conference on Scientific Information devoted an entire session to this problem.[13]

An enormous difficulty to be surmounted in designing any system for literature searching is that of preparation of the file to be searched. Every stage of the journey from natural language and figures, illustrations, and diagrams (the forms in which it is assumed most literature would be found initially), through the analysis by professional personnel, (to extract pertinent material, or to abstract or to transfer material _en toto_), to the final form of the information on some magnetic medium, is studded with formidable obstacles. Thyllis Williams's concern regarding the preservation of the informative content of a document is a very real one. It is extremely important, but difficult at this stage of development of most systems, to insure that there is adequacy of analysis of documents, and proper coverage of all material contained in them. It is perhaps for these reasons that Yngve insists so strongly on preservation of the natural language input as one prerequisite for proper subsequent retrieval.

The accuracy of any data which have been handled by humans at any point is suspect; on the other hand, machines which handle data cannot use "human judgment", and it is difficult, if not impossible, to formulate rules for processing every situation when diverse kinds of data are handled by machines. There simply does not exist any foolproof method at this time for transferring hard copy information directly to a magnetic medium with any assurance of obtaining an error-free file.

The basic approach to the design of a system for information retrieval varies widely. For some retrieval problems a set of descriptive terms called "descriptors" is stored for each entry in the file. Some retrieval attempts require

the specification of identical descriptors in order to recover the stored information. A broad or general specification might deluge the questioner with answers if (a) the file is extremely large or (b) the question is not asked in sufficient detail to screen out all of the unwanted items having duplicate descriptors. One criticism of this kind of file is its unwieldiness because of size; another is the difficulty of phrasing questions in order to avoid large numbers of "false drops". One fashion of forming a file for such a system is by means of an "inverted file" structure where, instead of listing under a document the descriptors applying to it, the file categories are made up of characteristics of material in the documents, and under each such descriptive heading is a list of the documents which contain material of such a category.

A more succinct file can be obtained by the use of precise definitions for wanted information from the file, but this procedure is not possible in very many applications. Chemical literature lends itself to representation in this fashion very readily, since it permits exact definition in terms of molecular formula and topological arrangement. Where an exact definition can be specified, any question which is formed for the purpose of retrieval must be equally exact if only pertinent responses are to be elicited.

At times, a more powerful search technique may lie in "screening" out the documents one does not want, in order to leave a residue of documents more likely to respond. By applying Boolean techniques to pre-stored symbols for descriptive material, it is sometimes possible to make lightning pre-searches of the file which shorten by several orders of magnitude the time which a serial search of the entire file would require.

Sometimes arrangement of the files in terms of the most frequently desired items is preferable, particularly if a serial procedure of searching is employed. The greatest activity is then concentrated in the area where the information most often required is clustered, and a more efficient operation results. Most searches, regardless of the manner of arrangement of material in the files, are serial in nature, according to the plan for retrieval. An ideal system would permit an inquiry to be made of the complete file simultaneously, with a simultaneous response from every entry which is being called for; at this time neither equipment nor systems for retrieval have reached this stage of development. An approach to it, however, is in the concept of selective use of random-access memories in terms of addressability, where symbols for information coincide with the address where the information is stored, so that instead of serial searching, selective searching is employed.

Sir Robert Watson-Watt has stated:

> "When he asks for a random-access memory, the client usually wants to have whatever may be the converse of a serial-access memory. Physically the answer has reduced to a very rapid-access (and still, most frequently, a serial-access) device, but philosophically, in the end it need not, and practically it should not, do so."

He suggests that the operation which is usually desired is a "Parallel partial-identifier search" and he further points out that what is wanted is "assured timely individual retrieval by selection, preferably by parallel interrogation of the system's 'memory' ".[14] Parallel rapid-access memories of the type which are needed for literature searching have been promised by design engineers; until the development of such memories has been completed, however, a simulated parallel memory can be partially realized by more sophisticated system designs which permit interrogation of only likely respondents by selective addressing or summary screening techniques. Henry Semarne has applied symbolic logic to screening techniques; [15] it is expected that this area will attract increasing numbers of workers as additional large-scale systems develop.

A large amount of effort has gone into automatic indexing and abstracting of English language text, because of the sheer magnitude of this task. In an effort to force computers to assist in this effort, H. P. Luhn, together with a group associated with him at IBM, has for several years been concerned with "auto-abstracting" of documents by computers, where the abstracted portion is based upon a statistical analysis of the frequency of occurrence of words, in context.[16] The abstracts are made up of whole sentences from the text of the document, where the sentences are chosen not only with respect to the presence of certain words, but also with respect to the relationship of those words to each other, in terms of their location in the sentence. Word lists are formed, excluding articles, prepositions, conjunctions, etc., and the most frequently used words are assumed to be of "high significance." From the number of these words in a sentence, and how closely they are clustered among the other words in the sentence, a "sentence significance factor" is computed. Those sentences with high scores based on these factors are extracted from the document to form an Auto-Abstract.

Some information retrieval systems have adroit indexing schemes for relating discrete portions of files when a unified response is desired. In this way, segments or fragments may be recovered in the context in which they are physically located, or they may be tied together in diverse ways for several levels of maneuvers. Complexity of such arrangements of course requires more intricate designs for the manipulation of data so arranged.

Another form of disposition of files is that proposed by Newell, Simon and Shaw, [17] where data making up an entity are not necessarily arranged together in contiguous memory locations; instead, the data form a "trail" where each datum carries the association with what precedes it and what succeeds it. This particular arrangement allows facility of amplification of data entries without massive rearrangements of data in memory. It also permits new entries or inserts, without dislocating former entries. In the same manner, it facilitates deletions of unwanted or no longer pertinent information. The unwanted locations are simply returned to storage for later use by successive entries. Unfortunately, with present equipment, this type of arrangement demands a certain amount of interpretive programming and requires a very large random access memory.

Information processing language (IPL) VI was employed to describe the operations defined in the paper cited above; the authors are now engaged in developing IPL VII, which is expected to be able to handle more complex situations.

Many literature searching systems do not take advantage of some of the advanced techniques with respect to data arrangement and format. The logical problems of setting up an information retrieval system for computers are knotty enough, without the application of several levels of sophistication in programming techniques and data arrangement. In addition, the incorporation of certain intellectual advancements to programs now in existence is probably not worthwhile until a plateau of development in literature searching is reached, because of their high cost in both time and money. Some of the improvements which could be put into programs now in existence might well be abandoned before they have proved themselves because of advances in the state of the art.

Although rapid advances continue to be made in the speed with which arithmetic calculations can be performed, computer manufacturers have been slow to incorporate into computers special devices which would aid materially in making searches. At least one computer manufacturer, however, is now incorporating "search units" into a general purpose computer. The inclusion of such equipment in computer design (perhaps as an optional feature) could well become a trend, and would not onlycut down on search time by a very large amount, but would make unnecessary some of the refinements in programming techniques which are costly, but which will become increasingly necessary in the absence of special hardware characteristics.

Certainly one of the difficulties encountered in all literature searching programs is the lack of appropriate automata with which to conduct the searches. Computers were originally designed to compute; there has been a lag in satisfying the need for automation in other areas where the need is just as great as for computers. Extremely large random-access memories with rapid "symbol manipulators" are required. Since those who are designing systems for literature searching do not in general have available the kind of automata which specifically meet their requirements, they must simulate such equipment by the use of computers. It is presumed that even this difficulty will be surmounted in the not too distant future, particularly since so much interest has been awakened in information retrieval in the last few years. Indeed, there are in existence some special purpose equipments for literature searching which can handle limited areas of the general problem at this time (although commercially available machines specifically designed as literature searching machines have not yet come into widespread use). A brief mention of some of them follows:

Several machines make use of photographic storage, some of which sense optically coded patterns:

1. The Rapid Selector, developed by Bush and Shaw, now under further development at the National Bureau of Standards for the Navy's Bureau of Ships.
2. FLIP, developed by the Benson-Lehner Corporation
3. Minicard, Eastman Kodak Company
4. Filmorex System, Filmorex, Paris (FRANCE)

In the following two machines, questions are entered by means of a special typewriter, and inquiries are made of a file, stored on magnetic tape. Several searches can be made simultaneously:

1. WRU Searching Selector, Western Reserve University
2. GE-250 (General Electric Information Searching Selector), a transistorized version of the WRU Searching Selector

Two machines made by IBM are the 9900 Special Index Analyzer and the 9310 Universal Card Scanner; each is a punched card machine.

It is hoped that from the brief discussion presented above, the reader will recognize the challenge inherent in the development of information retrieval systems, both from the standpoint of analysis and design of adequate systems to meet the challenge, and from the standpoint of design of suitable equipment to implement such systems.

BIBLIOGRAPHY

1. "ERMA -- Electronic Bookkeeper", Standford Research Institute News Bulletin, Volume 7, No. 9, October 1955; "Magnetic Reader Speeds Travelers-Check Processing", by K. R. Eldredge, et al, Control Engineering, Vol. 5, No. 7, July, 1958, p. 79-83.

2. "Operations Research and the Automation of Banking Procedures" by R. A. Byerly, Proceedings of the Fifth Annual Computer Applications Symposium, Armour Research Institute, Chicago, Ill., Oct. 29-30, 1958, p.1-10.

3. "Magnetic Ink Character Recognition--the Common Machine Language for Check Handling", Subcommittee Report, Banking--Journal of the American Bankers Association, Vol. 49, No. 2., August 1956.

4. Third Annual Progress Report of the Analysis and Development Program for Processing Letter-Size-Mail. National Bureau of Standards Report No. 6452 to the U. S. Post Office Department, by M. C. Stark and I. Rotkin, Dec. 31, 1958.

5. "Finding Chemical Records by Digital Computers" by Louis C. Ray and Russell A. Kirsch, Science, 126:3278 (25 October 1957) 814-819.

6. "A Survey of Automatic Reading Techniques" by M. E. Stevens, National Bureau of Standards, Washington, D. C. (In preparation).

7. "Machine Translation in the U.S.S.R.--A Brief Survey" by Sonya G. Machelson, American Documentation, Vol. XI, No. 2., April 1960, p. 151-154

8. "Documentation, Indexing, and Retrieval of Scientific Information". Report of the Committee on Government Operations, United States Senate, May 24, 1960. John L. McClellan (Ark.), Chairman.

10. "Problems in Scientific Communication" by Eric de Grolier, IBM Journal of Research and Development, 2:4 (October 1958) 276-281.

11. "A Programming Language for Mechanical Translation" by Victor H. Yngve, Mechanical Translation, 5:1 (July 1958) 24-41.

12. "A Program of Research and Development on Information Searching Systems" Summary. Itek Corp., Information Services Dept., Waltham, Mass. Oct. 1958. 7 p.; "A research Program on Information Searching Systems" by John W. Kuipers, Itek Corp., Waltham, Mass., Aug. 5, 1959; "Translating from Ordinary Discourse into Formal Logic--a Preliminary Systems Study. Scientific Report." by Thyllis Williams, ACF Industries, Inc., Alexandriz, Va., Sept. - Nov. 1956. Contract AF 19 (604) 1582. Tech. note No. AFCRC-TN-56-770. ASTIA doc. no. AD-98 813.

13. International Conference on Scientific Information, Washington, D. C., Nov. 16-21, 1958. (Proceedings.) Washington, National Academy of Sciences--National Research Council (AREA 6, Vol. 2., p. 1269-1409).

14. "Are Computers Important?" by Sir Robert Watson-Watt, Proceedings of the Eastern Joint Computer Conference, New York, N. Y., Dec. 10-12, 1956. p. 67-68.

15. "Symbolic Language in Language Engineering" by H. M. Semarne, Proceedings of the Western Joint Computer Conference, San Francisco, May 3-5, 1960, p. 61-71.

16. "The Preparation of Auto-Abstracts on the IBM 704 Data Processing System" by T. R. Savage, IBM Research Center, Yorktown Heights, Nov. 17, 1958.

17. "A Command Structure for Complex Information Processing" by J. C. Shaw, A. Newell, H. A. Simon and T. O. Ellis. Proceedings of the Western Joint Computer Conference, Los Angeles, May 6-8, 1958. p. 119-128.

U. S. DEPARTMENT OF AGRICULTURE GRADUATE SCHOOL
SEMINAR ON DATA PROCESSING FOR FEDERAL EXECUTIVES
SECOND SECTION

Selected References on Data Processing

1. Data Processing in Navy Management Information Systems, Sec. Nav. Inst. P10462.7. April 16, 1959.

2. Survey of Progress and Trend of Development and Use of ADP in Business and Management Control Systems of the Federal Government. June 1958. (This is a report to Congress by the Controller General)

3. Introduction to Data Processing. Haskins and Sells, 1957.

4. The Computer Challenge, E. D. Dwyer. Navy Management Review. August 59

5. Automatic Data Processing Equipment Policy and Methods Reports. John Diebold and Associates. Or

6. Office Automation and Office Automation Applications, by R. Hunt Brown. (Looseleaf handbooks with updating services) or

7. Data Processing Digest. A service of Canning, Sisson and Associates. Consists of a digest of periodicals.

8. General Service Contract Terms and Conditions with E.D.P. Manufacturers.

9. Electronic Data Processing for Business and Industry by Richard G. Canning. John Wiley and Sons, 1956.

10. Computers--Their Operation and Applications. By Edmund C. Berkeley and Lawrence Wainwright. Reinhold Publishing Company. 1956.

11. Data Processing by Electronics, by Haskins and Sells, 1955.

12. American Management Association Reports: Keeping Pace With Automation, Special Report No. 7; Electronic Data Processing in Industry, Special Report No. 3; Establishing an Integrated Data Processing System, Special Report No. 11; Electronics in Action, Special Report No. 22.

13. Use of Electronic Data Processing Equipment. A Congressional Hearing June 5, 1959.

14. An introduction to Automatic Computers, by Ned Chapin. (Princeton, New Jersey-195-7-D. Van Nostrand Co.)

15. Electronic Computers and Management Control. Kozmetsky, George and Paul Kircher. New York, McGraw-Hill Book Co., 1956.

16. Thinking Ahead: Potentials of Management Science; Peter F. Drucker; Harvard Business Review, Jan-Feb., 1959.

17. Electronics in Management. Lowell H. Hattery and George P. Bush. The University Press of Washington, D. C.

18. Installing Electronic Data Processing Systems; Richard G. Canning. John Wiley and Sons, Inc. 1957.

19. Concepts and Applications of Electronic Data Processing. U. S. Army Signal Corps

20. Eckert, W. J. and Jones, Rebecca, Faster, Faster, McGraw-Hill Book Company, Inc., New York, New York.

21. Livesley, R. K., An Introduction to Automatic Digital Computers, Cambridge University Press, New York, New York.

22. McCracken, D. D., Digital Computer Programming, John Wiley and Sons, Inc., New York, New York.

23. An Evaluation of Information Retrieval Systems, Memorandum Report No. 7170, September 30, 1959 -- a Report of Operations Research by Lockheed Aircraft Corporation, California Division

UNITED STATES DEPARTMENT OF AGRICULTURE

GRADUATE SCHOOL

WASHINGTON 25, D. C.

TO: Participants in Automatic Data Processing Seminar, Mar. 1960,

FROM: Seminar Coordinator

SUBJECT: Automatic Data Processing Talks

Here is your copy of another paper given as part of our seminar. We hope they will serve as an effective review.

Other papers will be sent to you as we receive them and get them reproduced.